

SEQUENCE LISTING

- 5 (1) GENERAL INFORMATION:
- (i) APPLICANT: Commonwealth Scientific and Industrial
Research Organisation and
Pacific Seeds Pty. Ltd.
- 10 (ia) INVENTORS: P. D. CHRISTIAN, K. H. J. GORDON and
T. N. HANZLIK
- (ii) TITLE OF INVENTION: INSECT VIRUSES AND THEIR USES IN
- 15 (iii) NUMBER OF SEQUENCES: 53
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: DAVIES COLLISON CAVE
(B) STREET: 1 LITTLE COLLINS STREET
(C) CITY: MELBOURNE
20 (D) STATE: VICTORIA
(E) COUNTRY: AUSTRALIA
(F) ZIP: 3000
- (v) COMPUTER READABLE FORM:
- 25 (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- 30 (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
(B) FILING DATE: 13 AUGUST 1993
(C) CLASSIFICATION:
- 35 (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: JOHN M. SLATTERY
(B) REGISTRATION NUMBER: NA
(C) REFERENCE/DOCKET NUMBER: 1613611
- 40 (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (613) 254 2777

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

5 (2) INFORMATION FOR SEQ ID NO:6:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 10 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
 15 CCATCGATGC CGGACTGGTA TCCCAGAGGA C 31

20 (2) INFORMATION FOR SEQ ID NO:7:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 25 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
 30 CCATCGATGA TCCAGCCTCC TCGCGGCGCC GGATGGGCA 39

35 (2) INFORMATION FOR SEQ ID NO:8:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 40 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
 45 GCTCTAGATC CATTGCGCAT CGAAGATGC CCATCCGGC 39

50 (2) INFORMATION FOR SEQ ID NO:9:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 55 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
 60 CCATCGATTG ATGCCGAGAA GGTAACCGA GAAACACAC 39

65 (2) INFORMATION FOR SEQ ID NO:10:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 70 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
 75 GCTCTAGACC AGGTAATATA CCACAACGTG TGTTCCTCT 39

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGGGGAATT CATTTAGGTG ACACATAGT TCTGCCTCCC CGGAC

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGGGGGATCC TGGTATCCCA GGGGGGGC

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCGGAAGCTT GTTTTCTTT CTTTACCA

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGGGGATCCG ATGGTATCCC GAGGGACGCT CAGCAGGTGG CATAGG

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAATAATTTT GTTACTTTAG AAGGAGATAT ACATATGAGC GAGCGAGCAC AC

(2) INFORMATION FOR SEQ ID NO:16:

85

75

75

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein - N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Gly Asp Ala Gly Val Ala Ser Gln
 1 5

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 6..32

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCCAT ATG AGC GAG GCC GGC GTC GCG TCA CAG 32
 Met Ser Glu Ala Gly Val Ala Ser Gln
 1 5

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein - N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Ser Glu Ala Gly Val Ala Ser Gln
 1 5

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATG GGA GAT GCT GGA GTG GCG TCA CAG 27
 Met Gly Asp Ala Gly Val Ala Ser Gln
 1 5

5

15

20

30

35

50

55

65

70

80

5

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

15

26

20

25

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

35

54

40

102

45

150

50

198

55

246

60

342

65

390

70

438

75

436

GCC TCC CGC ACC TTC TGC GTC GAC GGA GTC GGC TCT TGC GCG TTC AAA
Ala Ser Arg Thr Phe Cys Val Asp Gly Val Gly Ser Cys Ala Phe Lys
155 160 165

5	TCG CGC GTT GGA ATT GCC AAT CAC TCC CTC TAT GAC GTG ACC CTA GAG Ser Arg Val Gly Ile Ala Asn His Ser Leu Tyr Asp Val Thr Leu Glu 170 175 180	582
10	GAG CTG GCC AAT GCG TTT GAG AAC CAC GGA CTT CAC ATG GTC CGC GCG Glu Leu Ala Asn Ala Phe Glu Asn His Gly Leu His Met Val Arg Ala 185 190 195	630
15	TTC ATG CAC ATG CCA GAA GAG CTG CTC TAC ATG GAC AAC GTG GTT AAT Phe Met His Met Pro Glu Glu Leu Leu Tyr Met Asp Asn Val Val Asn 200 205 210	678
20	GCC GAG CTC GGC TAC CGC TTC CAC GTT ATT GAA GAG CCT ATG GCT GTG Ala Glu Leu Gly Tyr Arg Phe His Val Ile Glu Glu Pro Met Ala Val 215 220 225 230	726
25	AAG GAC TGC GCA TTC CAG GGG GGG GAC CTC CGT CTC CAC TTC CCT GAG Lys Asp Cys Ala Phe Gln Gly Gly Asp Leu Arg Leu His Phe Pro Glu 235 240 245	774
30	TTG GAC TTC ATC AAC GAG AGC CAA GAG CGG CSC ATC GAG AGG CTG GCC Leu Asp Phe Ile Asn Glu Ser Gln Glu Arg Arg Ile Glu Arg Leu Ala 250 255 260	822
35	GCC CGC GGC TCC TAC TCC AGA CGC GCC GTC ATT TTC TCC GGC GAC GAC Ala Arg Gly Ser Tyr Ser Arg Arg Ala Val Ile Phe Ser Gly Asp Asp 265 270 275	870
40	GAC TGG GGT GAT GCG TAC TTA CAC GAC TTC CAC ACA TGG CTC GCC TAC Asp Trp Gly Asp Ala Tyr Leu His Asp Phe His Thr Trp Leu Ala Tyr 280 285 290	918
45	CTA CTG GTG AGG AAC TAC CCC ACT CCG TTT GGT TTC TCA CTC CAT ATA Leu Leu Val Arg Asn Tyr Pro Thr Pro Phe Gly Phe Ser Leu His Ile 295 300 305 310	966
50	GAA GTC CAG AGG CGC CAC GGC TCC AGC ATT GAG CTG CGC ATC ACT CGC Glu Val Gln Arg Arg His Gly Ser Ser Ile Glu Leu Arg Ile Thr Arg 315 320 325	1014
55	GCG CCA CCT GGA GAC CGC ATG CTG GCC GTC GTC CCA AGG ACG TCC CAA Ala Pro Pro Gly Asp Arg Met Leu Ala Val Val Pro Arg Thr Ser Gln 330 335 340	1062
60	GGC CTC TGC AGA ATC CCA AAC ATC TTT TAT TAC GCC GAC GCG TCG GGC Gly Leu Cys Arg Ile Pro Asn Ile Phe Tyr Tyr Ala Asp Ala Ser Gly 345 350 355	1110
65	ACT GAG CAT AAG ACC ATC CTT ACG TCA CAG CAC AAA GTC AAC ATG CTG Thr Glu His Lys Thr Ile Leu Thr Ser Gln His Lys Val Asn Met Leu 360 365 370	1158
70	CTC AAT TTT ATG CAA ACG CGT CCT GAG AAG GAA CTA GTC GAC ATG ACC Leu Asn Phe Met Gln Thr Arg Pro Glu Lys Glu Leu Val Asp Met Thr 375 380 385 390	1206
75	GTC TTG ATG TCG TTC GCG CGC GCT AGG CTG CGC GCG ATC GTG GTC GCC Val Leu Met Ser Phe Ala Arg Ala Arg Leu Arg Ala Ile Val Val Ala 395 400 405	1254
80	TCA GAA GTC ACC GAG AGC TCC TGG AAC ATC TCA CCG GCT GAC CTG GTC Ser Glu Val Thr Glu Ser Ser Trp Asn Ile Ser Pro Ala Asp Leu Val 410 415 420	1302
85	CGC ACT GTC GTG TCT CTT TAC GTC CTC CAC ATC ATC GAG CGC CGA AGG Arg Thr Val Val Ser Leu Tyr Val Leu His Ile Ile Glu Arg Arg Arg 425 430 435	1350
90	GCT GCG GTC GCT GTC AAG ACC GCC AAG GAC GAC GTC TTT GGA GAG ACT Ala Ala Val Ala Val Lys Thr Ala Lys Asp Asp Val Phe Gly Glu Thr 440 445 450	1398
95	TCG TTC TGG GAG AGT CTC AAG CAC GTC TTG GGC TCC TGT TGC GGT CTG Ser Phe Trp Glu Ser Leu Lys His Val Leu Gly Ser Cys Cys Gly Leu 455 460 465 470	1446
100	CGC AAC CTC AAA GGC ACC GAC GTC GTC TTT ACT AAG CGC GTC GTC GAT Arg Asn Leu Lys Gly Thr Asp Val Val Phe Thr Lys Arg Val Val Asp 475 480 485	1494
105	AAG TAC CGA GTC CAC TCG CTC GGA GAC ATA ATC TGC GAC GTC CGC CTG Lys Tyr Arg Val His Ser Leu Gly Asp Ile Ile Cys Asp Val Arg Leu 490 495 500	1542

		TCC CCT GAA CAG GTC GGC TTC CTG CCG TCC CGC GTA CCA CCT GCC CGC	1590
		Ser Pro Glu Gln Val Gly Phe Leu Pro Ser Arg Val Pro Pro Ala Arg	
		505 510 515	
5		GTC TTT CAC GAC AGG GAA GAG CTT GAG GTC CTT CGC GAA GCT GGC TGC	1638
		Val Phe His Asp Arg Glu Leu Glu Val Leu Arg Glu Ala Gly Cys	
		520 525 530	
10		TAC AAC GAA CGT CCG GTA CCT TCC ACT CCT CCT GTG GAG GAG CCC CAA	1686
		Tyr Asn Glu Arg Pro Val Pro Ser Thr Pro Pro Val Glu Glu Pro Gln	
		535 540 545 550	
15		GGT TTC GAC GCC GAC TTG TGG CAC GCG ACC GCG GCC TCA CTC CCC GAG	1734
		Gly Phe Asp Ala Asp Leu Trp His Ala Thr Ala Ala Ser Leu Pro Glu	
		555 560 565	
20		TAC CGC GCC ACC TTG CAG GCA GGT CTC AAC ACC GAC GTC AAG CAG CTC	1782
		Tyr Arg Ala Thr Leu Gln Ala Gly Leu Asn Thr Asp Val Lys Gln Leu	
		570 575 580	
		AAG ATC ACC CTC GAG AAC GCC CTC AAG ACC ATC GAC GGG CTC ACC CTC	1830
		Lys Ile Thr Leu Glu Asn Ala Leu Lys Thr Ile Asp Gly Leu Trp Leu	
		585 590 595	
25		TCC CCA GTC AGA GGC CTC GAG ATG TAC GAG GGC CCG CCA GGC AGC GGC	1878
		Ser Pro Val Arg Gly Leu Glu Met Tyr Glu Gly Pro Pro Gly Ser Gly	
		600 605 610	
30		AAG ACG GGC ACC CTC ATC GCC GCC CTT GAG GCC GCG GGC GGT AAA GCA	1926
		Lys Thr Gly Thr Leu Ile Ala Ala Leu Glu Ala Ala Gly Gly Lys Ala	
		615 620 625 630	
35		CTT TAC GTG GCA CCC ACC AGA GAA CTG AGA GAG GCT ATG GAC CGG CGG	1974
		Leu Tyr Val Ala Pro Trp Arg Glu Leu Arg Glu Ala Met Asp Arg Arg	
		635 640 645	
40		ATC AAA CCG CCG TCC GCC TCG GCT ACG CAA CAT GTC GCC CTT GCG ATT	2022
		Ile Lys Pro Pro Ser Ala Ser Ala Thr Gln His Val Ala Leu Ala Ile	
		650 655 660	
		CTC CGT CGT GCC ACC GCC GAG GGC GCC CCT TTC GCT ACC GTG GTT ATC	2070
		Leu Arg Arg Ala Thr Ala Glu Gly Ala Pro Phe Ala Thr Val Val Ile	
		665 670 675	
45		GAC GAG TGC TTC ATG TTC CCG CTC GTG TAC GTC GCG ATC GTG CAC GCC	2118
		Asp Glu Cys Phe Met Phe Pro Leu Val Tyr Val Ala Ile Val His Ala	
		680 685 690	
50		TTG TCC CCG AGC TCA CGA ATA GTC CTT GTA GGG GAC GTC CAC CAA ATC	2166
		Leu Ser Pro Ser Ser Arg Ile Val Leu Val Gly Asp Val His Gln Ile	
		695 700 705 710	
55		GGG TTT ATA GAC TTC CAA GGC ACA AGC GCG AAC ATG CCG CTC GTT CGC	2214
		Gly Phe Ile Asp Phe Gln Gly Thr Ser Ala Asn Met Pro Leu Val Arg	
		715 720 725	
60		GAC GTC GTT AAG CAG TGC CGT CCG CGC ACT TTC AAC CAA ACC AAG CGC	2262
		Asp Val Val Lys Gln Cys Arg Arg Thr Phe Asn Gln Thr Lys Arg	
		730 735 740	
		IGT CCG GCC GAC GTC GTT GCC ACC ACG TTT TTC CAG AGC TTG TAC CCC	2310
		Cys Pro Ala Asp Val Val Ala Thr Thr Phe Phe Gln Ser Leu Tyr Pro	
		745 750 755	
65		GGG TGC ACA ACC ACC TCA GGG TGC GTC GCA TCC ATC AGC CAC GTC GCC	2358
		Gly Cys Thr Thr Thr Ser Gly Cys Val Ala Ser Ile Ser His Val Ala	
		760 765 770	
70		CCA GAC TAC CGC AAC AGC CAG GCG CAA ACG CTC TGC TTC ACG CAG GAG	2406
		Pro Asp Tyr Arg Asn Ser Gln Ala Gln Thr Leu Cys Phe Thr Gln Glu	
		775 780 785 790	
75		GAA AAG TCG CGC CAC GGG GCT GAG GGC GCG ATG ACT GTG CAC GAA GCG	2454
		Glu Lys Ser Arg His Gly Ala Glu Gly Ala Met Thr Val His Glu Ala	
		795 800 805	
		CAG GGA CGC ACT TTT GCG TCT GTC ATT CTG CAT TAC AAC GGC TCC ACA	2502
		Gln Gly Arg Thr Phe Ala Ser Val Ile Leu His Tyr Asn Gly Ser Thr	
		810 815 820	
80		GCA GAG CAG AAG CTC CTC GCT GAG AAG TCG CAC CTT CTA GTC GGC ATC	2550
		Ala Glu Gln Lys Leu Leu Ala Glu Lys Ser His Leu Leu Val Gly Ile	
		825 830 835	

20

5	GCC AAA ATC AAC GTA GAG CTT GCC GAA CCC GAC GCG ACC CCG AAG CCG Ala Lys Ile Asn Val Glu Leu Ala Glu Pro Asp Ala Thr Pro Lys Pro 935 940 945 950	2886
10	CAT AGG GCG TTC CAG GAA GGG GTA CAG TGG GTC AAG GTC ACC AAC GCG His Arg Ala Phe Gln Glu Gly Val Gln Trp Val Lys Val Thr Asn Ala 955 960 965	2934
15	TCT AAC AAA CAC CAG GCG CTC CAG ACG CTG TTG TCC GCG TAC ACC AAG Ser Asn Lys His Gln Ala Leu Gln Thr Leu Leu Ser Arg Tyr Thr Lys 970 975 980	2982
20	CGA AGC GCT GAC CTG CCG CTA CAC GAA GCT AAG GAG GAC GTC AAA CCG Arg Ser Ala Asp Leu Pro Leu His Glu Ala Lys Glu Asp Val Lys Arg 985 990 995	3030
25	ATG CTA AAC TCG CTT GAC CGA CAT TGG GAC TGG ACT GTC ACT GAA GAC Met Leu Asn Ser Leu Asp Arg His Trp Asp Trp Thr Val Thr Glu Asp 1000 1005 1010	3078
30	GCC CGT GAC CGA GCT GTC TTC GAG ACC CAG CTC AAG TTC ACC CAA CCG Ala Arg Asp Arg Ala Val Phe Glu Thr Gln Leu Lys Phe Thr Gln Arg 1015 1020 1025 1030	3126
35	GGC GGC ACC GTC GAA GAC CTG CTG GAG CCA GAC GAC CCC TAC ATC CGT Gly Gly Thr Val Glu Asp Leu Leu Glu Pro Asp Asp Pro Tyr Ile Arg 1035 1040 1045	3174
40	GAC ATA GAC TTC CTT ATG AAG ACT CAG CAG AAA GTG TCG CCC AAG CCG Asp Ile Asp Phe Leu Met Lys Thr Gln Gln Lys Val Ser Pro Lys Pro 1050 1055 1060	3222
45	ATC AAT ACG GGC AAG GTC GGG CAG GGG ATC GCC GCT CAC TCA AAG TCT Ile Asn Thr Gly Lys Val Gly Gln Gly Ile Ala Ala His Ser Lys Ser 1065 1070 1075	3270
50	CTC AAC TTC GTC CTC GCC GCT TGG ATA CCG ATA CTC GAG GAG ATA CTC Leu Asn Phe Val Leu Ala Ala Trp Ile Arg Ile Leu Glu Glu Ile Leu 1080 1085 1090	3318
55	CGT ACC GGG AGC CCG ACG GTC CCG TAC AGC AAC GGT CTC CCC GAC GAA Arg Thr Gly Ser Arg Thr Val Arg Tyr Ser Asn Gly Leu Pro Asp Glu 1095 1100 1105 1110	3366
60	GAA GAG GCC ATG CTG CTC GAA GCG AAG ATC AAT CAA GTC CCA CAC GCC Glu Glu Ala Met Leu Leu Glu Ala Lys Ile Asn Gln Val Pro His Ala 1115 1120 1125	3414
65	ACG TTC GTC TCG GCG GAC TGG ACC GAG TTT GAC ACC GCC CAC AAT AAC Thr Phe Val Ser Ala Asp Trp Thr Glu Phe Asp Thr Ala His Asn Asn 1130 1135 1140	3462
70	ACG AGT GAG CTG CTC TTC GCC GCC CTT TTA GAG CGC ATC GGC ACG CCT Thr Ser Glu Leu Leu Phe Ala Ala Leu Leu Glu Arg Ile Gly Thr Pro 1145 1150 1155	3510
75	GCA GCT GCC GTT AAT CTA TTC AGA GAA CGG TGT GGG AAA CCG ACC TTG Ala Ala Ala Val Asn Leu Phe Arg Glu Arg Cys Gly Lys Arg Thr Leu 1160 1165 1170	3558
80	CGA GCG AAG GGT CTA GGC TCC GTT GAA GTC GAC GGT CTG CTC GAC TCC Arg Ala Lys Gly Leu Gly Ser Val Glu Val Asp Gly Leu Leu Asp Ser 1175 1180 1185 1190	3606
85	GGC GCA GCT TGG ACG CCT TGC CGC AAC ACC ATC TTC TCT GCC GCC GTC Gly Ala Ala Trp Thr Pro Cys Arg Asn Thr Ile Phe Ser Ala Ala Val 1195 1200 1205	3654
90	ATG CTC ACG CTC TTC CGC GGC GTC AAG TTC GCA GCT TTC AAA GGC GAC Met Leu Thr Phe Arg Gly Val Lys Phe Ala Ala Phe Lys Gly Asp 1210 1215 1220	3702
95	GAC TCG CTC CTC TGT GGT AGC CAT TAC CTC CGT TTC GAC GCT AGC CCG Asp Ser Leu Leu Cys Gly Ser His Tyr Leu Arg Phe Asp Ala Ser Arg 1225 1230 1235	3750
100	CTT CAC ATG GGC GAA CGT TAC AAG ACC AAA CAT TTG AAG GTC GAG GTG Leu His Met Gly Glu Arg Tyr Lys Thr Lys His Leu Lys Val Glu Val 1240 1245 1250	3798
105	CAG AAA ATC GTG CCG TAC ATC GGA CTC CTC GTC TCC GCT GAG CAG GTC Gln Lys Ile Val Pro Tyr Ile Gly Leu Leu Val Ser Ala Glu Gln Val 1255 1260 1265 1270	3846

	GTC CTC GAC CCT GTC AGG AGC GCT CTC AAG ATA TTT GGG CGC TGC TAC Val Leu Asp Pro Val Arg Ser Ala Leu Lys Ile Phe Gly Arg Cys Tyr 1275 1280 1285	3894
5	ACA AGC GAA CTC CTT TAC TCC AAG TAC GTG GAG GCT GTG AGA GAC ATC Thr Ser Glu Leu Leu Tyr Ser Lys Tyr Val Glu Ala Val Arg Asp Ile 1290 1295 1300	3942
10	ACC AAG GGC TGG AGT GAC GCC CGC TAC CAC AGC CTC CTG TGC CAC ATG Thr Lys Gly Trp Ser Asp Ala Arg Tyr His Ser Leu Leu Cys His Met 1305 1310 1315	3990
15	TCA GCA TGC TAC TAC AAT TAC GCG CCG GAG TCT GCG GCG TAC ATC ATC Ser Ala Cys Tyr Tyr Asn Tyr Ala Pro Glu Ser Ala Ala Tyr Ile Ile 1320 1325 1330	4038
20	GAC GCT GTT GTT CGC TTT GGG CGC GGC GAC TTC CCG TTT GAA CAA CTG Asp Ala Val Val Arg Phe Gly Arg Gly Asp Phe Pro Phe Glu Gln Leu 1335 1340 1345 1350	4086
	CGC GTG GTG CGT GCC CAT GTG CAG GCA CCC GAC GCT TAC AGC AGC ACG Arg Val Val Arg Ala His Val Gln Ala Pro Asp Ala Tyr Ser Ser Thr 1355 1360 1365	4134
25	TAI CCG GCT AAC GTG CGC GCA TCG TGC CTT GAC CAC GTC TTC GAG CCC Tyr Pro Ala Asn Val Arg Ala Ser Cys Leu Asp His Val Phe Glu Pro 1370 1375 1380	4182
30	CGC CAG GCC GCC GCC CCG GCA GGT TTC GTT GCG ACA TGT GCG AAG CCG Arg Gln Ala Ala Ala Pro Ala Gly Phe Val Ala Thr Cys Ala Lys Pro 1385 1390 1395	4230
35	GAA ACG CCT TCT TCA CTT ACC GCG AAA GCT GGT GTT TCT GCG ACT ACA Glu Thr Pro Ser Ser Leu Thr Ala Lys Ala Gly Val Ser Ala Thr Thr 1400 1405 1410	4278
40	AGC CAC GTT GCG ACT GGG ACT GCG CCC CCG GAG TCT CCA TGG GAT GCA Ser His Val Ala Thr Gly Thr Ala Pro Pro Glu Ser Pro Trp Asp Ala 1415 1420 1425 1430	4326
	CCT GCA GCC AAC AGC TTT TCG GAG TTA TTG ACA CCG GAG ACC CCG TCC Pro Ala Ala Asn Ser Phe Ser Glu Leu Leu Thr Pro Glu Thr Pro Ser 1435 1440 1445	4374
45	ACA TCA TCC TCG CCG TCA TCG TCT TCA TCG GAC TCC TCT ACA TCG TGT Thr Ser Ser Ser Pro Ser Ser Ser Ser Asp Ser Ser Thr Ser Cys 1450 1455 1460	4422
50	GGA AGG TCG CTC AGT GGT GGA GAC ACC GCA AGG ACC ACA GAA GAC TTG Gly Arg Ser Leu Ser Gly Gly Asp Thr Ala Arg Thr Thr Glu Asp Leu 1465 1470 1475	4470
55	AAC AGC AGA AAG CCG CCT TCG CAA GAC AGG CAA TCA CGC TCG TCT GAA Asn Ser Arg Lys Pro Pro Ser Gln Asp Arg Gln Ser Arg Ser Ser Glu 1480 1485 1490	4518
60	TGT CTG GAC AGA AGC GGA GAA AGG ACA GGC AGT TCG TTA ACT GCC CCC Cys Leu Asp Arg Ser Gly Glu Arg Thr Gly Ser Ser Leu Thr Ala Pro 1495 1500 1505 1510	4566
	ACT GCT CCG AGC CCC TCA TTC TCA TTT TCG GAA AGA GCT CGA CTG GCG Thr Ala Pro Ser Pro Ser Phe Ser Phe Ser Glu Arg Ala Arg Leu Ala 1515 1520 1525	4614
65	ACC GGG CCG ACT GTC GCC GCT GCG ACA TCA CCT TCG GCA ACC CCA TCC Thr Gly Pro Thr Val Ala Ala Ala Thr Ser Pro Ser Ala Thr Pro Ser 1530 1535 1540	4662
70	TGC GCC ACG GAC CAG GTT GCC GCG AGG ACC ACG CCG GAC TTT GCG CCT Cys Ala Thr Asp Gln Val Ala Ala Arg Thr Thr Pro Asp Phe Ala Pro 1545 1550 1555	4710
75	TTC CTG GGT TCC CAG TCT GCC CGT GCT GTC TCG AAG CCG TAC CGG CCC Phe Leu Gly Ser Gln Ser Ala Arg Ala Val Ser Lys Pro Tyr Arg Pro 1560 1565 1570	4758
80	CCC ACG ACT GCC CGT TGG AAA GAA GTC ACC CCG CTC CAC GCG TGG AAG Pro Thr Thr Ala Arg Trp Lys Glu Val Thr Pro Leu His Ala Trp Lys 1575 1580 1585 1590	4806
	GGC GTG ACC GGA GAC CGA CCG GAA GTC AGG GAG GAC CCG GAG ACA GCG Gly Val Thr Gly Asp Arg Pro Glu Val Arg Glu Asp Pro Glu Thr Ala 1595 1600 1605	4854
85	GCG GTC GTC CAG GCT CTG ATC AGC GGC CGT TAT CCT CAG AAG ACG AAG	4902

	Ala	Val	Val	Gln	Ala	Leu	Ile	Ser	Gly	Arg	Tyr	Pro	Gln	Lys	Thr	Lys	
				1610					1615					1620			
5	CTT	TCC	TCC	GAC	GCA	TCC	AAA	GGC	TAC	TCA	AGA	ACT	AAG	GGA	TGC	TCA	4950
	Leu	Ser	Ser	Asp	Ala	Ser	Lys	Gly	Tyr	Ser	Arg	Thr	Lys	Gly	Cys	Ser	
				1625				1630					1635				
	CAA	TCC	ACC	TCT	TTT	CCT	GCC	CCG	AGT	GCG	GAT	TAC	CAG	GCC	CGC	GAC	4998
10	Gln	Ser	Thr	Ser	Phe	Pro	Ala	Pro	Ser	Ala	Asp	Tyr	Gln	Ala	Arg	Asp	
				1640			1645					1650					
	TGC	CAG	ACA	GTC	CGA	GTC	TGC	CGC	GCC	GCT	GCA	GAG	ATG	GCG	CGC	TCA	5046
	Cys	Gln	Thr	Val	Arg	Val	Cys	Arg	Ala	Ala	Ala	Glu	Met	Ala	Arg	Ser	
15				1655		1660					1665				1670		
	TGT	ATT	CAC	GAG	CCG	TTG	GCT	TCA	TCT	GCC	GCC	AGT	GCC	GAC	TTG	AAG	5094
	Cys	Ile	His	Glu	Pro	Ile	Ala	Ser	Ser	Ala	Ala	Ser	Ala	Asp	Leu	Lys	
				1675					1680					1685			
20	CGC	ATA	CGC	TCT	ACC	TCG	GAC	TCT	GTT	CCC	GAT	GTA	AAG	ATC	AGC	AAG	5142
	Arg	Ile	Arg	Ser	Thr	Ser	Asp	Ser	Val	Pro	Asp	Val	Lys	Ile	Ser	Lys	
				1690					1695				1700				

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

	Val	Pro	Arg	Thr	Ser	Gln	Gly	Leu	Cys	Arg	Ile	Pro	Asn	Ile	Phe	Tyr
				340										350		
5	Tyr	Ala	Asp	Ala	Ser	Gly	Thr	Glu	His	Lys	Thr	Ile	Leu	Thr	Ser	Gln
			355					360					365			
	His	Lys	Val	Asn	Met	Leu	Leu	Asn	Phe	Met	Gln	Thr	Arg	Pro	Glu	Lys
10			370				375					380				
	Glu	Leu	Val	Asp	Met	Thr	Val	Leu	Met	Ser	Phe	Ala	Arg	Ala	Arg	Leu
			385			390					395					400
15	Arg	Ala	Ile	Val	Val	Ala	Ser	Glu	Val	Thr	Glu	Ser	Ser	Trp	Asn	Ile
				405						410					415	
	Ser	Pro	Ala	Asp	Leu	Val	Arg	Thr	Val	Val	Ser	Leu	Tyr	Val	Leu	His
			420						425				430			
20	Ile	Ile	Glu	Arg	Arg	Arg	Ala	Ala	Val	Ala	Val	Lys	Thr	Ala	Lys	Asp
			435					440					445			
	Asp	Val	Phe	Gly	Glu	Thr	Ser	Phe	Trp	Glu	Ser	Leu	Lys	His	Val	Leu
25			450				455					460				
	Gly	Ser	Cys	Cys	Gly	Leu	Arg	Asn	Leu	Lys	Gly	Thr	Asp	Val	Val	Phe
			465			470					475					480
30	Thr	Lys	Arg	Val	Val	Asp	Lys	Tyr	Arg	Val	His	Ser	Leu	Gly	Asp	Ile
				485						490					495	
	Ile	Cys	Asp	Val	Arg	Leu	Ser	Pro	Glu	Gln	Val	Gly	Phe	Leu	Pro	Ser
			500						505					510		
35	Arg	Val	Pro	Pro	Ala	Arg	Val	Phe	His	Asp	Arg	Glu	Glu	Leu	Glu	Val
			515					520					525			
	Leu	Arg	Glu	Ala	Gly	Cys	Tyr	Asn	Glu	Arg	Pro	Val	Pro	Ser	Thr	Pro
40			530				535					540				
	Pro	Val	Glu	Glu	Pro	Gln	Gly	Phe	Asp	Ala	Asp	Leu	Trp	His	Ala	Thr
			545			550					555					560
45	Ala	Ala	Ser	Leu	Pro	Glu	Tyr	Arg	Ala	Thr	Leu	Gln	Ala	Gly	Leu	Asn
				565						570					575	
	Thr	Asp	Val	Lys	Gln	Leu	Lys	Ile	Thr	Leu	Glu	Asn	Ala	Leu	Lys	Thr
			580						585					590		
50	Ile	Asp	Gly	Leu	Thr	Leu	Ser	Pro	Val	Arg	Gly	Leu	Glu	Met	Tyr	Glu
			595					600					605			
	Gly	Pro	Pro	Gly	Ser	Gly	Lys	Thr	Gly	Thr	Leu	Ile	Ala	Ala	Leu	Glu
55			610				615					620				
	Ala	Ala	Gly	Gly	Lys	Ala	Leu	Tyr	Val	Ala	Pro	Thr	Arg	Glu	Leu	Arg
			625			630					635					640
60	Glu	Ala	Met	Asp	Arg	Arg	Ile	Lys	Pro	Pro	Ser	Ala	Ser	Ala	Thr	Gln
				645						650					655	
	His	Val	Ala	Leu	Ala	Ile	Leu	Arg	Arg	Ala	Thr	Ala	Glu	Gly	Ala	Pro
			660						665					670		

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

	Arg Phe Asp Ala Ser Arg Leu His Met Gly Glu Arg Tyr Lys Thr Lys	1235	1240	1245
5	His Ileu Lys Val Glu Val Gln Lys Ile Val Pro Tyr Ile Gly Leu Leu	1250	1255	1260
	Val Ser Ala Glu Gln Val Val Leu Asp Pro Val Arg Ser Ala Leu Lys	1265	1270	1275
10	Ile Phe Gly Arg Cys Tyr Thr Ser Glu Leu Leu Tyr Ser Lys Tyr Val	1285	1290	1295
	Glu Ala Val Arg Asp Ile Thr Lys Gly Trp Ser Asp Ala Arg Tyr His	1300	1305	1310
15	Ser Leu Leu Cys His Met Ser Ala Cys Tyr Tyr Asn Tyr Ala Pro Glu	1315	1320	1325
20	Ser Ala Ala Tyr Ile Ile Asp Ala Val Val Arg Phe Gly Arg Gly Asp	1330	1335	1340
	Phe Pro Phe Glu Gln Leu Arg Val Val Arg Ala His Val Gln Ala Pro	1345	1350	1355
25	Asp Ala Tyr Ser Ser Thr Tyr Pro Ala Asr Val Arg Ala Ser Cys Leu	1365	1370	1375
	Asp His Val Phe Glu Pro Arg Gln Ala Ala Ala Pro Ala Gly Phe Val	1380	1385	1390
30	Ala Thr Cys Ala Lys Pro Glu Thr Pro Ser Ser Leu Thr Ala Lys Ala	1395	1400	1405
35	Gly Val Ser Ala Thr Thr Ser His Val Ala Thr Gly Thr Ala Pro Pro	1410	1415	1420
	Glu Ser Pro Trp Asp Ala Pro Ala Ala Asn Ser Phe Ser Glu Leu Leu	1425	1430	1435
40	Thr Pro Glu Thr Pro Ser Thr Ser Ser Ser Pro Ser Ser Ser Ser	1445	1450	1455
	Asp Ser Ser Thr Ser Cys Gly Arg Ser Leu Ser Gly Gly Asp Thr Ala	1460	1465	1470
45	Arg Thr Thr Glu Asp Leu Asn Ser Arg Lys Pro Pro Ser Gln Asp Arg	1475	1480	1485
	Gln Ser Arg Ser Ser Glu Cys Leu Asp Arg Ser Gly Glu Arg Thr Gly	1490	1495	1500
50	Ser Ser Leu Thr Ala Pro Thr Ala Pro Ser Pro Ser Phe Ser Phe Ser	1505	1510	1515
55	Glu Arg Ala Arg Leu Ala Thr Gly Pro Thr Val Ala Ala Ala Thr Ser	1525	1530	1535
	Pro Ser Ala Thr Pro Ser Cys Ala Thr Asp Gln Val Ala Ala Arg Thr	1540	1545	1550
60	Thr Pro Asp Phe Ala Pro Phe Leu Gly Ser Gln Ser Ala Arg Ala Val	1555	1560	1565
65	Ser Lys Pro Tyr Arg Pro Pro Thr Thr Ala Arg Trp Lys Glu Val Thr	1570	1575	1580
	Pro Leu His Ala Trp Lys Gly Val Thr Gly Asp Arg Pro Glu Val Arg	1585	1590	1595
70	Glu Asp Pro Glu Thr Ala Ala Val Val Gln Ala Leu Ile Ser Gly Arg	1605	1610	1615
	Tyr Pro Gln Lys Thr Lys Leu Ser Ser Asp Ala Ser Lys Gly Tyr Ser	1620	1625	1630
75	Arg Thr Lys Gly Cys Ser Gln Ser Thr Ser Phe Pro Ala Pro Ser Ala	1635	1640	1645
80	Asp Tyr Gln Ala Arg Asp Cys Gln Thr Val Arg Val Cys Arg Ala Ala	1650	1655	1660
	Ala Glu Met Ala Arg Ser Cys Ile His Glu Pro Leu Ala Ser Ser Ala	1665	1670	1675

Ala Ser Ala Asp Leu Lys Arg Ile Arg Ser Thr Ser Asp Ser Val Pro
 1685 1690 1695

5 Asp Val Lys Ile Ser Lys Ser Ala
 1700

10 (2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5312 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA

20 (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 4218..4512

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

25 GTTCTGCCTC CCCCGGACGG TAAATATAGG GGAACAATGT ACGCGAAAGC GACAGACGTG 60
 GCGCGTGTCT ACGCCGCGGC AGATGTCGCC TACGCGAACG TACTGCAGCA GAGAGCAGTC 120
 30 AAGTTGGA CTGCGCTTCA AAGGGGGCAG TTIACCCCGC ACACAACACC CGATCCTGGC CGGGCACCAC 240
 CTGCTCGCAG AAGAGGTTCT GCACAATTTC GCCAGGGGAC GTAGCACAGT GCTCGAGATA 300
 35 GGGCCGTCTC TGCACAGCGC ACTTAAGCTA CATGGGGCAC CGAACGCCCC C3TCGCAGAC 360
 TATCAGGGGT GCACCAAGTA CGGCACCCGC GACGGCTCGC GACACATTAC GGCCTTAGAG 420
 40 TCTAGATCCG TCGCCACAGG CCGGCCCGAG TTCAAGGCCG ACGCCTCACT GCTCGCCAAC 480
 GGCATTGCCT CCCGCACCTT CTGCGTCGAC GGAGTCGGCT CTGCGCGTT CAAATCGCGC 540
 45 GTTGGAATTG CCAATCACTC CCTCTATGAC GTGACCCTAG AGGAGCTGGC CAATGCGTTT 600
 GAGAACCACG GACTTCACAT GGTCCGCGCG TTCATGCACA TGCCAGAAGA GCTGCTCTAC 660
 ATGGACAACG TGTTAATGCG CGAGCTCGGC TACCGCTTCC ACGTATTGA AGAGCCTATG 720
 50 GCTGTGAAGG ACTGCGCATT CCAGGGGGGG GACCTCCGTC TCCACTTCCC TGAGTTGGAC 780
 TTCATCAACG AGAGCCAAGA GCGGCGCATC GAGAGGCTGG CCGCCCGCGG CTCCTACTCC 840
 55 AGACGCGCCG TCATTTTCTC CGGCGACGAC GACTGGGGTG ATGCGTACTT ACACGACTTC 900
 CACACATGGC TCGCCTACCT ACTGGTGAGG AACTACCCCA CTCGGTTTGG TTTCTCACTC 960
 CATATAGAAG TCCAGAGGCG CCACGGCTCC AGCATTGAGC TGCGCATCAC TCGCGCGCCA 1020
 60 CCTGGAGACC GCATGCTGGC CGTCGTCCCA AGGACGTCCC AAGGCCTCTG CAGAATCCCA 1080
 AACATCTTTT ATTACGCCGA CGCGTCGGGC ACTGAGCATA AGACCATCCT TACGTCACAG 1140
 65 CACAAAGTCA ACATGCTGCT CAATTTTATG CAAACGCGTC CTGAGAAGGA ACTAGTCGAC 1200
 ATGACCGTCT TGATGTCGTT CGCGCGCGCT AGGCTGCGCG CGATCGTGGT CGCCTCAGAA 1260
 GTCACCGAGA GCTCCTGGAA CATCTACCGG GCTGACCTGG TCCGCACTGT CGTGTCTCTT 1320
 70 TACGTCTCTC ACATCATCGA GCGCCGAAGG GCTGCGGTCT CTGTCAAGAC CGCCAAGGAC 1380
 GACGTCTTTG GAGAGACTTC GTTCTGGGAG AGTCTCAAGC ACGTCTTGGG CTCCTGTTGC 1440
 75 GGTCTGCGCA ACCTCAAAGG CACCGACGTC GTCTTTACTA AGCGCGTCGT CGATAAGTAC 1500
 CGAGTCCACT CGCTCGGAGA CATAATCTGC GACGTCGCCG TGTCCCCTGA ACAGGTCGGC 1560
 TTCTGCGCGT CCCGCGTACC ACCTGCCCGC GTCTTTACAG ACAGGGAAGA GCTTGAGGTC 1620
 80 CTTGCGGAAG CTGGCTGCTA CAACGAACGT CCGGTACCTT CCACTCCTCC TGTGGAGGAG 1680
 CCCCAGGTT TCGACGCCGA CTTGTGGCAC GCGACCGCGG CCTCACTCCC CGAGTACCGC 1740
 GCCACCTTGC AGGCAGGTCT CAACACCGAC GTCAAGCAGC TCAAGATCAC CCTCGAGAAC 1800

	GCCCTCAAGA CCATCGACGG GCTCACCCCTC TCCCCAGTCA GAGGCCTCGA GATGTACGAG	1860
	GGCCCGCCAG GCAGCGGCAA GACGGGCACC CTCATCGCCG CCCTTGAGGC CGCGGGCGGT	1920
5	AAAGCACTTT ACGTGGCACC CACCAGAGAA CTGAGAGAGG CTATGGACCG GCGGATCAAA	1980
	CCGCCGTCCG CCTCGGCTAC GCAACATGTC GCCCTTGCGA TTCTCCGTCG TGCCACCGCC	2040
10	GAGGGCGCCC CTTCGCTAC CGTGGTTATC GACGAGTGT TCATGITCCC GTCGTGTAC	2100
	GTCGCGATCG TGACAGCCTT GTCCCCGAGC TCACGAATAG TCCTTGATAGG GGACGTCCAC	2160
	CAAATCGGGT TTATAGACTT CCAAGGCACA AGCGCGAACA TGCCGCTCGT TCGCGACGTC	2220
15	GTTAAGCAGT GCCGTGGGG CACTTTCAAC CAAACCAAGC GCTGTCCGGC CGACGTGCTT	2280
	GCCACCACGT TTTTCCAGAG CTTGTACCCC GGGTGCACAA CCACCTCAGG GTGCGTCGCA	2340
20	TCCATCAGCC ACGTCGCCCC AGACTACCGC AACAGCCAGG CGCAAACGCT CTGCTTCACG	2400
	CAGGAGGAAA AGTCGCGCCA CGGGGCTGAG GCGCGCATGA CTGTGCACGA AGCGCAGGGA	2460
	CGCACTTTTG CGTCTGTCAT TCTGCATTAC AACGGCTCCA CAGCAGAGCA GAAGCTCCTC	2520
25	GCTGAGAAGT CGCACCTTCT AGTCGGCATC ACGCGCCACA CCAACCACCT GTACATCCGC	2580
	GACCCGACAG GTGACATTGA GAGACAATC AACCATAGCG CGAAAGCCGA GGTGTTTACA	2640
30	GACATCCCTG CACCCCTGGA GATCAGACT GTCAAACCGA GTGAAGAGGT GCAGCGCAAC	2700
	GAAGTGATGG CAACGATACC CCCGAGAGT GCCACGCCG ACGGAGCAAT CCATCTGCTC	2760
	CGCAAGAAGT TCGGGGACCA ACCCGACTGT GCGTGTGTCG CTTTGTCGAA GACCGGCTAC	2820
35	GABGIGITTG GCGGTCGTGT CAAATCAAC GTAGAGCTTG CCGAACCCGA CGCGACCCCG	2880
	AAGCCGATA GGGCCTTCCA GGAAGGGGTA CAGTGGTCA AGGTACACAA CGCGTCTAAC	2940
40	AAACACCAGG CGCTCCAGAC GCTGTTGTCC CGCTACACCA AGCGAAGCGC TGACCTGCCG	3000
	CTACACGAAG CTAAGGAGGA CGTCAAACGC ATGCTAAACT CGCTTGACCG ACATTGGGAC	3060
	TGGAAGTGA CTGAAGAGC CCGTGACCGA GCTGTCTTCG AGACCCAGCT CAAGTTCACC	3120
45	CAACGCGGCG GCACCGTCGA AGACCTGCTG GAGCCAGACG ACCCCTACAT CCGTGACATA	3180
	GACTTCCTTA TGAAGACTCA GCAGAAAGTG TCGCCCAAGC CGATCAATAC GGGCAAGGTC	3240
50	GGGCAGGGGA TCGCCGCTCA CTCAAAGTCT CTCAACTTCG TCCTCGCCCG TTGGATACGC	3300
	ATACTCGAGG AGATACTCCG TACCGGGAGC CGCACGGTCC GGTACAGCAA CGGTCTCCCC	3360
	GACGAAGAAG AGGCCATGCT GCTCGAAGCG AAGATCAATC AAGTCCACAA CGCCACGTTT	3420
55	GTCTCGGGGG ACTGGACCGA GTTTGACACC GCCCACAATA ACACGAGTGA GCTGCTCTTC	3480
	GCCGCCCTTT TAGAGCGCAT CGGCACGCCT GCAGCTGCCG TTAATCTATT CAGAGAACGG	3540
60	TGTGGGAAAC GCACCTTGCG AGCGAAGGGT CTAGGCTCCG TTGAAGTCGA CGGTCTGCTC	3600
	GACTCCGGCG CAGCTTGGAC GCCTTGCCGC AACACCATCT TCTCTGCCG CGTCATGCTC	3660
	ACGCTCTTCC GCGGCGTCAA GTTCGAGCT TTCAAAGGCG ACGACTCGCT CCTCTGTGGT	3720
65	AGCCATTACC TCCGTTTCGA CGCTAGCCGC CTCACATGG GCGAACGTTA CAAGACCAAA	3780
	CATTTGAAGG TCGAGGTGCA GAAATCGTG CCGTACATCG GACTCCTCGT CTCGCTGAG	3840
70	CAGGTCGTCC TCGACCTGT CAGGAGCGCT CTCGAATAT TTGGGCGCTG CTACACAAGC	3900
	GAACCTCCTT ACTCCAAGTA CGTGGAGGCT GTGAGAGACA TCACCAAGGG CTGGAGTGAC	3960
	GCCCGCTACC ACAGCCTCCT GTGCCACATG TCAGCATGCT ACTACAATTA CGCGCCGGAG	4020
75	TCTGCGCGCT ACATCATCGA CGCTGTGTTT CGCTTGGGC GCGGCGACTT CCCGTTTGAA	4080
	CAACTGCGCG TGGTGCCTGC CCATGTGCAG GCACCCGACG CTTACAGCAG CACGTATCCG	4140
80	GCTAACGTGC GCGCATCGTG CTTTGACCAC GTCTTCGAGC CCCGCCAGGC CGCGCCCCG	4200
	GCAGGTTTCG TTGCAGC ATG TGC GAA GCC GGA AAC GCC TTC TTC ACT TAC	4250
	Met Cys Glu Ala Gly Asn Ala Phe Phe Thr Tyr	
	1 5 10	

	CGC GAA AGC TGG TGT TTC TGC GAC TAC AAG CCA CGT TGC GAC TGG GAC	4298
	Arg Glu Ser Trp Cys Phe Cys Asp Tyr Lys Pro Arg Cys Asp Trp Asp	
	15 20 25	
5	TGC GCC CCC GGA GTC TCC ATG GGA TGC ACC TGC AGC CAA CAG CTT TTC	4346
	Cys Ala Pro Gly Val Ser Met Gly Cys Thr Cys Ser Gln Gln Leu Phe	
	30 35 40	
10	GGA GTT ATT GAC ACC GGA GAC CCC GTC CAC ATC ATC CTC GCC GTC ATC	4394
	Gly Val Ile Asp Thr Gly Asp Pro Val His Ile Ile Leu Ala Val Ile	
	45 50 55	
15	GTC TTC ATC GGA CTC CTC TAC ATC GTG TGG AAG GTC GCT CAG TGG TGG	4442
	Val Phe Ile Gly Leu Leu Tyr Ile Val Trp Lys Val Ala Gln Trp Trp	
	60 65 70 75	
20	AGA CAC CGC AAG GAC CAC AGA AGA CTT GAA CAG CAG AAA GCC GCC TTC	4490
	Arg His Arg Lys Asp His Arg Arg Leu Glu Gln Gln Lys Ala Ala Phe	
	80 85 90	
25	GCA AGA CAG GCA ATC ACG CTC GTC TGAATGTC TGGACAGAAG CCGAGAAAAGG	4542
	Ala Arg Gln Ala Ile Thr Leu Val	
	95	
30	ACAGGCAGTT CGTTAACTGC CCCACATGCT CCGAGCCCOCT CAITCTCATI TTCGAAAGA	4602
	GCTCGACTGG CGACCGGGCC GACTGTGCGC GCTGCGACAT CACCTTCGGC AACCCATCC	4662
35	TGCGCCACGG ACCAGGTTGC CGCGAGGACC ACGCCGGACT TTGCGCCTTT CCTGGGTTCC	4722
	CAGTCTGCCC GTGCTGTCTC GAAGCCGTAC CGGCCCCCA CGACTGCCCG TTGAAAGAA	4782
40	GTCACCCCGC TCCACGCGTG GAAGGGCGTG ACCGGAGACC GACCGGAAGT CAGGGAGGAC	4842
45	CCGGAGACAG CGGCGGTCTG CCAGGCCTG ATCAGCGGCC GTTATCCTCA GAAGACGAAG	4902
	CTTTCCTCCG ACGCAICCAA AGGCTACTCA AGAACTAAGG GATGCTCACA ATCCACCTCI	4962
50	TTTCCTGCCC CGAGTGCGGA TTACCAGGCC CGCGACTGCC AGACAGTCCG AGTCTGCCGC	5022
	GCCGCTGCAG AGATGGCGCG CTCATGTATT CACGAGCCGT TGGCTTCATC TGCCGCCAGT	5082
	GCCGACTTGA AGCGCATACG CTCTACCTCG GACTCTGTTC CCGAIGTAAA GATCAGCAAG	5142
	AGCGCATGAA GGAACAAAAT TAGTTTCCTT GTTCGTAAAC AAGGTGGTCC CTCOCATTGA	5202
	GGTAAAGACT CTGGTGAGTC CTCAACGTTA CTCGTIGAGT CTGCTGCGGT TCGATTCCAT	5262
	TCCAAGCAG CAAAGGGTGC GCAACTAGTA CGCGCCCCC TGGGATACCA	5312

5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

35
(2) INFORMATION FOR SEQ ID NO:43:

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

55	GTTCCTGCCTC CCCCGGACGG TAAATATAGG GGAACAAATGT ACGCGAAAGC GACAGACGTG	60
	GCGCGTGTCT ACGCCGCGGC AGATGTCGCC TACGCGAAGC TACTGCAGCA GAGAGCAGTC	120
60	AAGTTGGACT TCGCCCCGCC ACTGAAGGCA CTAGAAACCC TCCACAGACT GIACATATCCG	180
	CTGCGCTTCA AAGGGGGCAC TTTACCCCGC ACACAACACC CGATCTTGGC CGGGCACCAA	240
	CGTGTGCGAG AAGAGTTCT GCACAATTTC GCCAGGGGAC GTAGCACAGT GCTCGAGATA	300
65	GGGCCGTCTC TGCACAGCGC ACTTAAGCTA CATGGGGCAC CGAACGCCCC CGTCGCGAGC	360
	TATCACGGGT GCACCAAGTA CGGCACCCGC GACGGCTCGC GACACATTAC GGCCTTAGAG	420
70	TCTAGATCCG TCGCCACAGG CCGGCCCGAG TTCAAGGCCG ACGCCTCACT GCTCGCCAAC	480
	GGCATTGGCT CCCGCACCTT CTGCGCTGAC GGAGTCGGCT CTTGCGCGTT CAAATCGCGC	540
	GTTGGAATTG CCAATCACTC CCTCTATGAC GTGACCCTAG AGGAGCTGGC CAATGCGTTT	600
75	GAGAACCACG GACTTCACAT GGTCCGCGCG TTCATGCACA TGCCAGAAGA GCTGCTCTAC	660
	ATGGACAACG TGGTTAATGC CGAGCTCGGC TACCGCTTCC ACGTTATTGA AGAGCCTATG	720
80	GCTGTGAAGG ACTGCGCATT CCAGGGGGGG GACCTCCGTC TCCACTTCCC TGAGTTGGAC	780
	TTTCATCAACG AGAGCCAAGA GCGGCGCATC GAGAGGCTGG CCGCCGCGG CTCCTACTCC	840
	AGACGCGCCG TCATTTTCTC CGGCGACGAC GACTGGGGTG ATGCGTACTT ACACGACTTC	900

	CACACATGGC	TCGCCTACCT	ACTGGTGGAG	AATACCCCA	CTCCGTTTGG	TTTCTCACTC	960
	CATATAGAAG	TCCAGAGGCG	CCACGGCTCC	AGCATTGAGC	TGCGCATCAC	TCGCGCGCCA	1020
5	CCTGGAGACC	GCATGCTGGC	CGTCGTCCCA	AGGACGTCCC	AAGGCCTCTG	CAGAATCCCA	1080
	AACATCTTTT	ATTACGCCGA	CGCGTCGGGC	ACTGAGCATA	AGACCATCCT	TACGTCACAG	1140
10	CACAAAGTCA	ACATGCTGCT	CAATTTTATG	CAAACGCGTC	CTGAGAAGGA	ACTAGTCGAC	1200
	ATGACCGTCT	TGATGTCGTT	CGCGCGCGCT	AGGCTGCGCG	CGATCGTGGT	CGCCTCAGAA	1260
	GTCACCGAGA	GCTCCTGGAA	CATCTCACCG	GCTGACCTGG	TCGCGACTGT	CGTGTCTCTT	1320
15	TACGTCTCTC	ACATCATCGA	GCGCCGAAGG	GCTGCGGTGC	CTGTCAAGAC	CGCCAAGGAC	1380
	GACGTCTTTG	GAGAGACTTC	GTCTCGGGAG	AGTCTCAAGC	ACGICTTTGG	CTCCTGTTGC	1440
20	GGTCTGCGCA	ACCTCAAAGG	CACCGACGTC	GTCTTTACTA	AGCGCGTCGT	CGATAAGTAC	1500
	CGAGTCCACT	CGCTCGGAGA	CATAATCTGC	GACGTCCGCT	TGTCCTCTGA	ACAGGTCGGC	1560
	TTCTTGCCGT	CCCGCGTACC	ACCTGCCCCG	GTCTTTCACG	ACAGGGAAGA	GCTTGAGGTC	1620
25	CTTCGCGAAG	CTGCGCTGTA	CAACGAACGT	CCGGTACCTT	CCACTCCTCC	TGTGGAGGAG	1680
	CCCCAAGGTT	TCGACGCCGA	CTTGTCGGAC	GCGACCGGGG	CCTCACTCCC	CGAGTACCGC	1740
30	GCCACCTTGC	AGGCAGGTCT	CAACACCGAC	GTCAAGCAGC	TCAAGATCAC	CCTCGAGAAC	1800
	GCCCTCAAGA	CCATCGACGG	GCTCACCTCT	TCCCCAGTCA	GAGGCCTCGA	GATGTACGAG	1860
	GGCCCCCCAG	GCAGCGGCAA	GACGGGCACC	CTCATCGCCG	CCCTTGAGGC	CGCGGGCGGT	1920
35	AAAGCACTTI	ACGTGGCACC	CACCAGAGAA	CTGAGAGAGG	CTATG3ACCG	GCGGATCAAA	1980
	CCGCGGTCCG	CCTCGGCTAC	GCAACATGTC	GCCCTTGCGA	TICTCCGTCG	TGCCACCGCC	2040
40	GAGGGCGCCC	CTTTCCCTAC	CGTGCTTATC	GACGAGTGCT	TCA1GTTCCC	GCTCGTGTAC	2100
	GTGCGCATCG	TGCACGCCTT	GTCCCCGAGC	TCACGAATAG	TCCITGTAGG	GGACGTCCAC	2160
	CAAAATCGGGT	TTATAGACTT	CAAAGGCACA	AGCGCGAACA	TGCCGCTCTG	TCGCGACGTC	2220
45	GTTAAGCAGT	GCCGTCGGCG	CACCTTCAAC	CAAAACCAAGC	GCTGTCCGGC	CGACGTCGTT	2280
	GCCACCACGT	TTTTCCAGAG	CTTGATCCCC	GGGTGCACAA	CCACCTCAGG	GTGCGTCGCA	2340
50	TCCATCAGCC	ACGTGCCCCC	AGACTACCGC	AACAGCCAGG	CGCAAACGCT	CTGCTTCACG	2400
	CAGGAGGAAA	AGTCGCGCCA	CGGGGCTGAG	GGCGCGATGA	CTGTGCACGA	AGCGCAGGGA	2460
	CGCACTTTTG	CGTCTGTCAT	TCTGCATTAC	AACG3CTCCA	CAGCAGAGCA	GAAGCTCCTC	2520
55	GCTGAGAAGT	CGCACCTTCT	AGTCGGCATC	ACGCGCCACA	CCAACCACCT	GTACATCCGC	2580
	GACCCGACAG	GTGACATTGA	GAGACAACCT	AACCATAGCG	CGAAAGCCGA	GGTGTTTACA	2640
60	GACATCCCTG	CACCCCTGGA	GATCAGGACT	GTCAAACCGA	GTGAAGAGGT	GACAGCGAAC	2700
	GAAGTGATGG	CAACGATACC	CCCGCAGAGT	GCCACGCCGC	ACGGAGCAAT	CCATCTGCTC	2760
	CGCAAGAACT	TCGGGGACCA	ACCCGACTGT	GGCTGTGTCG	CTTTGGCGAA	GACCGGCTAC	2820
65	GAGGTGTTTG	GCGGTCGTGC	CAAAATCAAC	GTAGAGCTTG	CCGAACCCGA	GCGGACCCCG	2880
	AAGCCGCATA	GGGCGTTCCA	GGAAGGGGTA	CAGTGGGGTA	AGGTACACCA	CGCGTCTAAC	2940
70	AAACACCAGG	CGCTCCAGAC	GCTGTTGTCC	CGCTACACCA	AGCGAAGCGC	TGACCTGCCG	3000
	CTACACGAAG	CTAAGGAGGA	CGTCAAACGC	ATGCTAAACT	CGCTTGACCG	ACATTGGGAC	3060
	TGGACTGTCA	CTGAAGACGC	CCGTGACCGA	GCTGTCTTCG	AGACCCAGCT	CAAGTTCACC	3120
75	CAACGCGGCG	GCACCGTCGA	AGACCTGCTG	GAGCCAGACG	ACCCCTACAT	CCGTGACATA	3180
	GACTTCCCTA	TGAAGACTCA	CGAGAAAGTG	TCGCCCAAGC	CGATCAATAC	GGGCAAGGTC	3240
80	GGGCAGGGGA	TCGCCGCTCA	CTCAAAGTCT	CTCAACTTCG	TCCTCGCCGC	TTGGATACGC	3300
	ATACTCGAGG	AGATACTCCG	TACCGGGAGC	CGCACGGTCC	GGTACAGCAA	CGGTCTCCCC	3360
	GACGAAGAAG	AGGCCATGCT	GCTCGAAGCG	AAGATCAATC	AAGTCCCACA	GCCACGTTTC	3420

	GTCTCGGCGG ACTGGACCGA GTTTGACACC GCCACAATA ACACGAGTGA GCTGCTCTTC	3480
	GCCGCCCTTT TAGAGCGCAT CGGCACGCCT GCAGCTGCCG TTAATCTATT CAGAGAACGG	3540
5	TGTGGGAAAC GCACCTTGCG AGCGAAGGGT CTAGGCTCCG TTGAAGTCGA CGGTCTGCTC	3600
	GACTCCGCGC CAGCTTGGAC GCCTTGCCGC AACACCATCT TCTCTGCCGC CGTCATGCTC	3660
10	ACGCTCTTCC GCGGCGTCAA GTTCGCAGCT TTCAAAGGCG ACGACTCGCT CCTCTGTGGT	3720
	AGCCATTACC TCCGTTTCGA CGCTAGCCGC CTTACATGG GCGAACGTTA CAAGACCAAA	3780
	CATTTGAAGG TCGAGGTGCA GAAAATCGTG CCGTACATCG GACTCCTCGT CTCGCTGAG	3840
15	CAGGTCGTCC TCGACCTGT CAGGAGCGCT CTCGAAGATAT TTGGGCGCTG CTACACAAGC	3900
	GAACTCCTTT ACTCCAAGTA CGTGGAGGCT GTGAGAGACA TCACCAAGGG CTGGAGTGAC	3960
20	GCCCGCTACC ACAGCCTCCT GTGCCACATG TCAGCATGCT ACTACAATTA CGCGCCGGAG	4020
	TCTGCGCGCT ACAICATCGA CGCTGTGTT CGCTTTGGGC GCGGCGACTT CCCGTTIGAA	4080
	CAACTGCGCG TGGTGCCTGC CCATGTGCAG GCACCCGACG CTTACAGCAG CACGTATCCG	4140
25	GCTAACGTGC GCGCATCGTG CCTTGACCAC GTCTTCGAGC CCCGCCAGGC CGCGCCCCCG	4200
	GCAGGTTTCG TIGCGACATG TCGGAAGCCG GAAACGCCTT CTTCACTTAC CGCGAAAGCT	4260
30	GGTGTTTCTG CCACTACAAG CCACGTTGCG ACTGGGACTG CGCCCCCGGA GTCTCCATGG	4320
	GATGCACCTG CAGCCAACAG CTTTTCGGAG TTATTGACAC CGGAGACCCC GTCCACATCA	4380
	TCTTCGCCGT CATCGCTTTC ATCGGACTCC TCTACATCGT GTGGAAGGTC GCTCAGTGGT	4440
35	GGAGACACCG CAAGGACCAC AGAAGACTTG AACAGCAGAA AGCCGCTTTC GCAAGACAGG	4500
	CAATCACGCT CGTCTGA ATG TCT GGA CAG AAG CGG AGA AAG GAC AGG CAG	4550
40	Met Ser Gly Gln Lys Arg Arg Lys Asp Arg Gln	
	1 5 10	
	TTC GTT AAC IGC CCC CAC TGC TCC GAG CCC CTC ATT CTC ATT TTC GGA	4590
	Phe Val Asn Cys Pro His Cys Ser Glu Pro Leu Ile Leu Ile Phe Gly	
	15 20 25	
45	AAG AGC TCG ACT GGC GAC CGG GCC GAC TGT CGC CGC TGC GAC ATC ACC	4646
	Lys Ser Ser Thr Gly Asp Arg Ala Asp Cys Arg Arg Cys Asp Ile Thr	
	30 35 40	
50	TTC GGC AAC CCC ATC CTG CGC CAC GGA CCA GGT TGC CGC GAG GAC CAC	4694
	Phe Gly Asn Pro Ile Leu Arg His Gly Pro Gly Cys Arg Glu Asp His	
	45 50 55	
55	GCC GGA CTT TGC GCC TTT CCT GGG TTC CCA GTC TGC CCG TGC TGT CTC	4742
	Ala Gly Leu Cys Ala Phe Pro Gly Phe Pro Val Cys Pro Cys Cys Leu	
	60 65 70 75	
	GAA GCC GTA CCG GCC CCC CAC GAC TGC CCG TTG GAA AGA AGT CAC CCC	4790
	Glu Ala Val Pro Ala Pro His Asp Cys Pro Leu Glu Arg Ser His Pro	
	80 85 90	
60	GCT CCA CGC GTG GAA GGG CGT GAC CGG AGA CCG ACC GGA AGT CAG GGA	4838
	Ala Pro Arg Val Glu Gly Arg Asp Arg Arg Pro Thr Gly Ser Gln Gly	
	95 100 105	
65	GGA CCC GGA GAC AGC GGC GGT CGT CCA GGC TCT GAT CAG CGG CCG TTA	4886
	Gly Pro Gly Asp Ser Gly Gly Arg Pro Gly Ser Asp Gln Arg Pro Leu	
	110 115 120	
70	TCC TCA GAA GAC GAA GCT TTC CTC CGA CGC ATC CAA AGG CTA CTC AAG	4934
	Ser Ser Glu Asp Glu Ala Phe Leu Arg Arg Ile Gln Arg Leu Leu Lys	
	125 130 135	
75	AAC TAAGGGATGC TCACAATCCA CCTCTTTTCC TGCCCCGAGT GCGGATTACC	4987
	Asn	
	140	
	AGGCCCGCGA CTGCCAGACA GTCCGAGTCT GCCGCGCCGC TGCAGAGATG GCGCGCTCAT	5047
80	GTATTACGCA GCGGTTGGCT TCATCTGCCG CCAGTGCCGA CTTGAAGCGC ATACGCTCTA	5107
	CCTCGGACTC TGTTCCCGAT GTAAAGATCA GCAAGAGCGC ATGAAGGAAC AAAATTAGTT	5167
	TCCTTGITCG TAAACAAGGT GGTCCCTCCC ATTGAGGTAA AGACTCTGGT GAGTCCTCAA	5227

CGTTACTCGT TGAGTCTGCT GCGGTTTCGAT TCCATTCCCA AGCAGCAAAG GGTGCGCAAC 5287
TAGTACGGCG CCCCCTGGGA TACCA 5312

5

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Met Ser Gly Gln Lys Arg Arg Lys Asp Arg Gln Phe Val Asn Cys Pro
1 5 10 15
His Cys Ser Glu Pro Leu Ile Leu Ile Phe Gly Lys Ser Ser Thr Gly
20 25 30
Asp Arg Ala Asp Cys Arg Arg Cys Asp Ile Thr Phe Gly Asn Pro Ile
35 40 45
Leu Arg His Gly Pro Gly Cys Arg Glu Asp His Ala Gly Leu Cys Ala
50 55 60
Phe Pro Gly Phe Pro Val Cys Pro Cys Cys Leu Glu Ala Val Pro Ala
65 70 75 80
Pro His Asp Cys Pro Leu Glu Arg Ser His Pro Ala Pro Arg Val Glu
85 90 95
Gly Arg Asp Arg Arg Pro Thr Gly Ser Gln Gly Gly Pro Gly Asp Ser
100 105 110
Gly Gly Arg Pro Gly Ser Asp Gln Arg Pro Leu Ser Ser Glu Asp Glu
115 120 125
Ala Phe Leu Arg Arg Ile Gln Arg Leu Leu Lys Asn
130 135 140

45

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5312 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4944..5162

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

65 GTTCTGCCTC CCCCGGACGG TAAATATAGG GGAACAATGT ACGCGAAAGC GACAGACGTG 60
GCGCGTGTCT ACGCCGCGGC AGATGTGCGC TACGCGAACG TACTGCAGCA GAGAGCAGTC 120
70 AAGTTGGACT TCGCCCCGCC ACTGAAGGCA CTAGAAACCC TCCACAGACT GTACTATCCG 180
CTGCGCTTCA AAGGGGGCAC TTACCCCCG ACACAACACC CGATCCTGGC CGGGCACCAA 240
CGTGTGCGAG AAGAGGTTCT GCACAATTTC GCCAGGGGAC GTAGCACAGT GCTCGAGATA 300
75 GGGCCGTCTC TGCACAGCGC ACTTAAGCTA CATGGGGCAC CGAACGCCCC CGTCGCGAGC 360
TATCACGGGT GCACCAAGTA CGGCACCCGC GACGGCTCGC GACACATTAC GGCCTTAGAG 420
TCTAGATCCG TCGCCACAGG CCGGCCCGAG TTCAAGGCCG ACGCCTCACT GCTCGCCAAC 480
80 GGCATTGCCT CCCGCACCTT CTGCGTCGAC GGAGTCG3CT CTTGCGCGTT CAAATCGCGC 540
GTTGGAATTG CCAATCACTC CCTCTATGAC GTGACCCTAG AGGAGCTGGC CAATGCGTTT 600

	GAGAACCACG GACTTCACAT GGTCCGCGCG TTCATGCACA TGCCAGAAGA GCTGCTCTAC	660
	ATGGACAACG TGGTTAATGC CGAGCTCGGC TACCGCTTCC ACGTTATTGA AGAGCCTATG	720
5	GCTGTGAAGG ACTGCGCATT CCAGGGGGGG GACCTCCGTC TCCACTTCCC TGAGTTGGAC	780
	ITCATCAACG AGAGCCAAGA GCGGCGCATC GAGAGGCTGG CCGCCCGCGG CTCCTACTCC	840
10	AGACGCGCCG TCATTTTCTC CGGCGACGAC GACTGGGGTG ATGCGTACTT ACACGACTTC	900
	CACACATGGC TCGCCTACCT ACTGGTGAGG AACTACCCCA CTCGTTTGG TITCTCACTC	960
	CATATAGAAG TCCAGAGGCG CCACGGCTCC AGCATTGAGC TCGCATCAC TCGCGCGCCA	1020
15	CCTGGAGACC GCATGCTGGC CGTCGTCCCA AGGACGTCCC AAGCCCTCTG CAGAATCCCA	1080
	AACATCTTTT ATTACGCCGA CGCGTCGGGC ACTGAGCATA AGACCATCCT TACGTCACAG	1140
20	CACAAAGTCA ACATGCTGCT CAATTTTATG CAAACGCGTC CTGAGAAGGA ACTAGTCGAC	1200
	ATGACCGTCT TGATGTCTGT CGCGCGCGCT AGGCTGCGCG CGATCGTGGT CGCCTCAGAA	1260
	GTCACCGAGA GCTCCTGGAA CATCTCACCG GCTGACCTGG TCCGCACTGT CGTGTCTCTT	1320
25	TACGTCTTCC ACATCATCGA GCGCCGAAGG GCTGCGGTGG CTGTCAAGAC CGCCAAGGAC	1380
	GACGCTTTTG GAGAGACTTC GTTCTGGGAG AGTCTCAAGC ACGTCTTGGG CTCCTGTTGC	1440
30	GGTCTGCGCA ACCTCAAAGG CACCGACGTC GTCTTTACTA AGCGCGTCGT CGATAAGTAC	1500
	CGAGTCCACT CGCTCGGAGA CATAATCTGC GACGTCCGCC TGTCCCCTGA ACAGGTCGGC	1560
	TTCTTGCCGT CCCGCGTACC ACCTGCCCGC GTCTTTACAG ACAGGAAGA GCTTGAGGTC	1620
35	CTTCGCGAAG CTGGCTGCTA CAACGAACGT CCGGTACCTT CCACTCCTCC TGTGGAGGAG	1680
	CCCCAAGGTT TCGACGCCGA CTTGTGGCAC GCGACCGCGG CCTCACTCCC CGAGTACCGC	1740
40	GCCACCTTGC AGGCAGGTCT CAACACCGAC GTCAAGCAGC TCAAGATCAC CCTCGAGAAC	1800
	GCCCTCAAGA CCATCGACGG GCTCACCTC TCCCCAGTCA GAGCCCTCGA GATGTACGAG	1860
	GGCCCGCCAG GCAGCGGCAA GACGGGCACC CTCATCGCCG CCCTTGAGGC CGCGGGCGGT	1920
45	AAAGCACTTT ACGTGGCACC CACCAGAGAA CTGAGAGAGG CTATGGACCG GCGGATCAAA	1980
	CCGCGCTCCG CCTCGGTAC GCAACATGTC GGCCTTGCGA TTCTCCGTGG TGCCACCGCC	2040
50	GAGGGCGCCC CTTTCGCTAC CGTGGTTATC GACGAGTGCT TCATGTTCCT GCTCGTGTAC	2100
	GTCCGATCG TGACAGCCTT GTCCCGAGC TCACGAATAG TCCTTGATGG GGACGTCCAC	2160
	CAATCGGGT TTATAGACTT CCAAGGCACA AGCGCAACA TGCCGCTCGT TCGCGACGTC	2220
55	GTAAAGCAGT GCCGTGCGCG CACTTTCAAC CAAACCAAGC GCTGTCCGGC CGACGTCTGT	2280
	GCCACCACGT TTTTCCAGAG CTTGTACCCC GGGTGCAAA CCACCTCAGG GTGGGTGCA	2340
60	TCCATCAGCC ACGTCGCCCC AGACTACCGC AACAGCCAGG CGCAAACGCT CTGCTTCACG	2400
	CAGGAGGAAA AGTCGCGCCA CGGGGCTGAG GCGCGATGA CTGTGCACGA AGCGCAGGGA	2460
	CGCACTTTTG CGTCTGTCAT TCTGCATTAC AACGGCTCCA CAGCAGACGA GAAGCTCCTC	2520
65	GCTGAGAAGT CGCACCTTCT AGTCGGCATC ACGCGCCACA CCAACCACCT GTACATCCGC	2580
	GACCCGACAG GTGACATTGA GAGACAACCT AACCATAGCG CGAAAGCCGA GGTGTTTACA	2640
70	GACATCCCTG CACCCCTGGA GATCAGACT GTCAAACCGA GTGAAGAGGT GCAGCGCAAC	2700
	GAAGTGATGG CAACGATACC CCCGAGAGT GCCACGCGC ACGGAGCAAT CCATCTGCTC	2760
	CGCAAGAACT TCGGGGACCA ACCCGACTGT GGCTGTGTCT CTTTGGCGAA GACCGGCTAC	2820
75	GAGGTGTTTG GCGGTCGTGC CAAAATCAAC GTAGAGCTTG CCGAACCCTG CGCGACCCCG	2880
	AAGCCGCATA GGGCGTTCCA GGAAGGGGTA CAGTGGGTCA AGGTCACCAA CGCGTCTAAC	2940
80	AAACACCAGG CGCTCCAGAC GCTGTTGTCC CGCTACACCA AGCGAAGCGC TGACCTGCCG	3000
	CTACACGAAG CTAAGGAGGA CGTCAAACGC ATGCTAAACT CGCTTGACCG ACATTGGGAC	3060
	TGGACTGICA CTGAAGACGC CCGTGACCGA GCTGTCTTCG AGACCCAGCT CAAGTTCACC	3120

	CAACGCGGCG GCACCGTCGA AGACCTGCTG GAGCCAGACG ACCCCTACAT CCGTGACATA	3180
	GACTTCCTTA TGAAGACTCA GCAGAAAGTG TCGCCCAAGC CGATCAATAC GGGCAAGGTC	3240
5	GGGCAGGGGA TCGCCGCTCA CTCAAAGTCT CTCAACTTCG TCCTCGCCGC TTGGATACGC	3300
	ATACTCGAGG AGATACTCCG TACCGGGAGC CGCACGGTCC GGTACAGCAA CGGTCTCCCC	3360
10	GACGAAGAAG AGGCCATGCT GCTCGAAGCG AAGATCAATC AAGTCCACACA CGCCACGTTC	3420
	GTCTCGGCGG ACTGGACCGA GTTTGACACC GCCACAATA ACACGAGTGA GCTGCTCTTC	3480
	GCCGCCCTTT TAGAGCGCAT CGGCACGCCT GCAGCTGCCG TTAATCTATT CAGAGAACGG	3540
15	TGTGGGAAAC GCACCTTGCG AGCGAAGGGT CTAGGCTCCG TTGAAGTCGA CGGTCTGCTC	3600
	GACTCCGGCG CAGCTTGGAC GCCTTGCCGC AACACCATCT TCTCTGCCGC CGTCATGCTC	3660
20	ACGCTCTTCC GCGGCGTCAA GTTCGCAGCT TTCAAAGGCG ACGACTCGCT CCTCTGTGGT	3720
	AGCCATTACC TCCGTTTCGA CGCTAGCCGC CTTACATGG GCGAACGTTA CAAGACCAAA	3780
	CATTTGAAGG TCGAGGTGCA GAAATCGTG CCGTACATCG GACTCCCTCG CTCCGCTGAG	3840
25	CAGGTCGTCC TCGACCCPGT CAGGAGCGCT CTCAGATAT TTGGGCGGTG CTACACAAGC	3900
	GAACTCCTII ACTCCAAGTA CGTGGAGGCT GTGAGAGACA TCACCAAGGG CTGGAGTGAC	3960
30	GCCCGCTACC ACAGCCTCCT GTGCCACATG TCAGCATGCT ACTACAATTA CGCGCCGGAG	4020
	TCGCGGCGT ACATCATCGA CGCTGTGTGT CGCTTTGGGC GCGGCGACTT CCCGTTTGAA	4080
	CAACTGCGCG TGGTGGCTGC CCATGTGCAG GCACCCGACG CTTACAGCAG CACGTATCCG	4140
35	GCTAACGTGC GCGCATCGTG CTTTGACCAC GTCTTCGAGC CCCGCCAGGC CGCCGCCCCG	4200
	GCAGGTTTCG TTGCGACATG TGCGAAGCCG GAAACGCCTT CTTCACTTAC CGCGAAAGCT	4260
40	GGTGTTCCTG CGACTACAAG CCACGTTGCG ACTGGGACTG CGCCCCCGGA GTCTCCATGG	4320
	GATGCACCTG CAGCCAACAG CTTTTCGGAG TTATTGACAC CGGAGACCCC GTCCACATCA	4380
	TCCTCGCCGT CATCGTCTTC ATCGGACTCC TCTACATCGT GTGGAAGGTC GCTCAGTGGT	4440
45	GGAGACCCG CAAGGACCAC AGAAGACTTG AACAGCAGAA AGCCGCCTTC GCAAGACAGG	4500
	CAATCACGCT CGTCTGAATG TCTGGACAGA AGCGGAGAAA GGACAGGCAG TTCGTTAACT	4560
50	GCCCCCACTG CTCCGAGCCC CTCATTCTCA TTTTCGGAAA GAGCTCGACT GGCAGCCGGG	4620
	CCGACTGTGC CCGCTGCGAC ATCACCTTCG GCAACCCCAI CCTGCCCCAC GGACCAGGTT	4680
	GCCGCGAGGA CCACGCGGGA CTTTGCGCCT TTCTGGGTT CCCAGTCTGC CCGTGTCTGC	4740
55	TCGAAGCCGT ACCGGCCCCC CACGACTGCC CGTTGGAAG AAGTCACCCC GCTCCACGCG	4800
	TGGAAGGGCG TGACCGGAGA CCGACCGGAA GTCAGGGAGG ACCCGGAGAC AGCGGCGGTC	4860
60	GTCCAGGCTC TGATCAGCGG CCGTTATCCT CAGAAGCGA AGCTTTCCTC CGACGCATCC	4920
	AAAGGCTACT CAAGAACTAA GGG ATG CTC ACA ATC CAC CTC TTT TCC TGC	4970
	Met Leu Thr Ile His Leu Phe Ser Cys	
	1 5	
65	CCC GAG TGC GGA TTA CCA GGC CCG CGA CTG CCA GAC AGT CCG AGT CTG	5018
	Pro Glu Cys Gly Leu Pro Gly Pro Arg Leu Pro Asp Ser Pro Ser Leu	
	10 15 20 25	
70	CCG CGC CGC TGC AGA GAT GGC GCG CTC ATG TAT TCA CGA GCC GTT GGC	5066
	Pro Arg Arg Cys Arg Asp Gly Ala Leu Met Tyr Ser Arg Ala Val Gly	
	30 35 40	
75	TTC ATC TGC CGC CAG TGC CGA CTT GAA GCG CAT ACG CTC TAC CTC GGA	5114
	Phe Ile Cys Arg Gln Cys Arg Leu Glu Ala His Thr Leu Tyr Leu Gly	
	45 50 55	
	CTC TGT TCC CGA TGT AAA GAT CAG CAA GAG CGC ATG AAG GAA CAA AAT	5162
	Leu Cys Ser Arg Cys Lys Asp Gln Glu Arg Met Lys Glu Gln Asn	
	60 65 70	
80	TAGTTTCCTT GTTCGTAAAC AAGGTGGTCC CTCCATTGA GGTAAAGACT CTGGTGAGTC	5222
	CTCAACGTTA CTCGTTGAGT CTGCTGCGGT TCGATTCCAT TCCCAAGCAG CAAAGGGTGC	5282
85	GCAACTAGTA CGGCGCCCCC TGGGATACCA	5312

5

10

15

Met Leu Thr Ile His Leu Phe Ser Cys Pro Glu Cys Gly Leu Pro Gly
1 5 10 15

20

25

30

35

40

45

50

CCCGGCCCTT ACACACCTCG AGTCCGTGAC CACCGGATTA TACGTCGCCC ACCACACGGC 120

55

CCGCCTTAGC TGCACCTACG GCAGCGTTGA TAGCGCGGAT TT ATG AGC GAG CAC 294

60

Met Ser Glu His
1

Thr Ile Ala His Ser Ile Thr Leu Pro Pro Gly Tyr Thr Leu Ala Leu
5 10 15 20

65

Ile Pro Pro Glu Pro Glu Ala Gly Trp Glu Met Leu Glu Trp Arg His
25 30 35

70

Ser Asp Leu Thr Thr Val Ala Glu Pro Val Thr Phe Gly Ser Ala Pro

75

ACA CCG TCA CCG TCA ATG GTA GAA GAA ACC AAC GGC GTC GGA CCG GAA * 486
Thr Pro Ser Pro Ser Met Val Glu Glu Thr Asn Gly Val Gly Pro Glu
55 60 65

5	GGC AAG TTT CTC CCC CTG ACA ATT TCA CCG CTG CTG CAC AAG ACC TCG Gly Lys Phe Leu Pro Leu Thr Ile Ser Pro Leu Leu His Lys Thr Ser 70 75 80	534
10	CGC AAA GCC TTG ACG CCA ACA CCG TCA CTT TCC CCG CTA ACA TCT CTA Arg Lys Ala Leu Thr Pro Thr Pro Ser Leu Ser Pro Leu Thr Ser Leu 85 90 95 100	582
15	GCA TGC CCG AAT TCC GGA ATT GGG CCA AGG GAA AGA TCG ACC TCG ACT Ala Cys Pro Asn Ser Gly Ile Gly Pro Arg Glu Arg Ser Thr Ser Thr 105 110 115	630
20	CCG ATT CCA TCG GCT GGT ACT TCA AGT ACC TTG ACC CAG CGG GTG CTA Pro Ile Pro Ser Ala Gly Thr Ser Ser Thr Leu Thr Gln Arg Val Leu 120 125 130	678
25	CAG AGT CTG CGC GCG CCG TCG GCG AGI ACT CGA AGA TCC CTG ACG GCC Gln Ser Leu Arg Ala Pro Ser Ala Ser Thr Arg Arg Ser Leu Thr Ala 135 140 145	726
30	TCG TCA AGT TCT CCG TCG ACG CAG AGA TAAGAGAGAT CTATAACGAG Ser Ser Ser Pro Ser Thr Gln Arg 150 155	773
35	GAGTCCCCCG TCGTCACTGA CGTGTCGGTC CCCCTCGACG GCCGCCAGTG GAGCCTCTCG ATTTTCTCCT TTCCGATGTT CAGAACCGCC TACGTGCGCG TAGCGAACGT CGAGAACAAG	833
40	GAGATGTCGC TCGACGTGT CAACGACCTC ATCGAGTGGC TCAACAATCT CGCGACTGG CGTTATGTCG TTGACTCTGA ACAGTGGATT AACTTCACCA ATGACACCAC GTACTACGTC	893
45	CGCATCCGCG TTCTACGTCC AACCTACGAC GTTCCAGACC CCACAGAGGG CTTGTTCGC ACAGTCTCAG ACTACCGCCT CACTTATAAG GCGATAACAT GTGAAGCCAA CATGCCAACA	953
50	CTCGTCGACC AAGGCTTTTG GATCGGCGGC CAGTACGCTC TCACCCCGAC TAGCCTACCG CAGTACGACG TCAGCGAGGC CTACGCTCTG CACACTTTGA CCTTCGCCAG ACCATCCAGC	1013
55	GCCGCTGCAC TCGCGTTTGT GTGGGCAGGT TTGCCACAGG GTGGCACTGC GCCTGCAGGC ACTCCAGCCT GGGAGCAGGC ATCCTCGGGT GGCTACCTCA CCTGGCGCCA CAACGGTACT	1073
60	ACTTTCCAG CTGGCTCCGT TAGCTACGTT CTCCTGAGG GTTTCGCCCT TGAGCGCTAC GACCCGAACG ACGGCTCTTG GACCGACTTC GCTTCCGAG GAGACACCGT CACTTTCCGG	1133
65	CAGGTGCGCG TCGACGAGGT CGTTGTGACC AACAAACCCG CCGGCGGCGG CAGCGCCCCC ACCTTCACCG TGAGAGTGCC CCCTTCAAAC GCTTACACCA ACACCGTGT TAGGAACACG	1193
70	CTCTTAGAGA CTCGACCTC CTCTCGTAGG CTCGAACTCC CTATGCCACC TGCTGACTTI GGACAGACGG TCGCCAACAA CCCGAAGATC GAGCAGTCGC TTCTTAAAGA AACACTTGGC	1253
75	TGCTATTGG TCCACTCCAA AATGCGAAAC CCCGTTTCC AGCTCACGCC AGCCAGCTCC TTTGGCGCGG TTCTCTCAA CAATCCGGGT TATGAGCGCA CACGCGACCT CCCGACTAC	1313
80	ACTGGCATCC GTGACTCATT CGACCAGAAC ATGTCCACCG CTGTGGCCCA CTTCGCTCA CTCTCCCACT CCTGCAGTAT CGTCACTAAG ACCTACCAGG GTTGGGAAGG CGTCACGAAC	1373
	GTCAACACGC CTTTCGGCCA ATTGCGCAC GCGGCGCTCC TCAAGAATGA GGAGATCCTC TGCTTCGCGG ACGACCTGGC CACCCGTCTC ACAGGTGTCT ACCCGCCAC TGACAACTTC	1433
	GCGGCCCGG TTTCTGCCTT CGCCGCGAAC ATGCTGTCTT CCGTGCTGAA GTCGGAGGCA ACGTCTCCA TCATCAAGTC CGTTGGCGAG ACTGCGGTG GCGCGGCTCA GTCCGGCTC	1493
	GCGAAGCTAC CCGGACTGCT AATGAGTGT CAGGGAAGA TTGCCGCGG GTGCCGCGG CGCCGAGCGC GCGCGCGCGC CGCTCGTGCC AATTAGTTG CTCGCTCCTG TTTCCCGGTT	1553
	TCGTAAAACG GCGTGGTCCC GCACATTACG CGTACCCTAA AGACTCTGGT GAGTCCCCGT CGTTACACGA CCGGTCTGCC GCGGTTGAT TCCATTCCCA AGCGGCAAGA AGGACGTAGT	1613
	TAGCTCTGCG TCCCTCGGGA TACCA	1673
		2478

(2) INFORMATION FOR SEQ ID NO:48:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 157 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

15 Met Ser Glu His Thr Ile Ala His Ser Ile Thr Leu Pro Pro Gly Tyr
 1 5 10 15
 Thr Leu Ala Leu Ile Pro Pro Glu Pro Glu Ala Gly Trp Glu Met Leu
 20 20 25 30
 Glu Trp Arg His Ser Asp Leu Thr Thr Val Ala Glu Pro Val Thr Phe
 35 40 45
 Gly Ser Ala Pro Thr Pro Ser Pro Ser Met Val Glu Glu Thr Asn Gly
 50 55 60
 Val Gly Pro Glu Gly Lys Phe Leu Pro Leu Thr Ile Ser Pro Leu Leu
 65 70 75 80
 His Lys Thr Ser Arg Lys Ala Leu Thr Pro Thr Pro Ser Leu Ser Pro
 35 90 95
 Leu Thr Ser Leu Ala Cys Pro Asn Ser Gly Ile Gly Pro Arg Glu Arg
 100 105 110
 Ser Thr Ser Thr Pro Ile Pro Ser Ala Gly Thr Ser Ser Thr Leu Thr
 115 120 125
 Gln Arg Val Leu Gln Ser Leu Arg Ala Pro Ser Ala Ser Thr Arg Arg
 130 135 140
 Ser Leu Thr Ala Ser Ser Ser Ser Pro Ser Thr Gln Arg
 145 150 155

(2) INFORMATION FOR SEQ ID NO:49:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2478 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 366..2306

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

65 GTTTTCTTT CTTTACCAAG TGTGGTAAAA TTAAACAAA GAAGAAAACC AGGACCGTAA 60
 CCGGGCCCTT ACACACCTCG AGTCCGTGAC CACCGGATTA TACGTCGCCC ACCACACGGC 120
 GCCTTTTCCG ACCACTCTCG AGAGTCGTTG GGAGTTTCGT CCGTGACCAC CCGGTTGGCA 180
 70 GTCGACAGAC GCTTCCGGAC CACTAGAACC TCCTCGAGCG ACGCACACAC AGCACACACA 240
 CCGCCTTAGC TGCACCTACG GCAGCGTTGA TAGCGCGGAT TTATGAGCGA GCACACCATC 300
 75 GCCCACTCCA TCACATTACC ACCCGGTTAC ACCCTTGCCC TAATACCCCC TGAACCTGAA 360
 GCAGG ATG GGA GAT GCT GGA GTG GCG TCA CAG CGA CCT CAC AAC CGT 407
 Met Gly Asp Ala Gly Val Ala Ser Gln Arg Pro His Asn Arg
 1 5 10
 80 CGC GGA ACC CGT AAC GTT CGG GTC AGC GCC AAC ACC GTC ACC GTC AAT 455
 Arg Gly Thr Arg Asn Val Arg Val Ser Ala Asn Thr Val Thr Val Asn
 15 20 25 30
 85 GGT AGA AGA AAC CAA CGG CST CGG ACC GGA AGG CAA GTT TCT CCC CCT 503
 Gly Arg Arg Asn Gln Arg Arg Arg Thr Gly Arg Gln Val Ser Pro Pro

5

5	AAT TGG GCC AAG GGA AAG ATC GAC CTC GAC TCC GAT TCC ATC GGC TGG Asn Trp Ala Lys Gly Lys Ile Asp Leu Asp Ser Asp Ser Ile Gly Trp 80 85 90	647
10	TAC TTC AAG TAC CTT GAC CCA GCG GGT GCT ACA GAG TCT GCG CGC GCC Tyr Phe Lys Tyr Leu Asp Pro Ala Gly Ala Thr Glu Ser Ala Arg Ala 95 100 105 110	695
15	GTC GGC GAG TAC TCG AAG ATC CCT GAC GGC CTC GTC AAG TTC TCC GTC Val Gly Glu Tyr Ser Lys Ile Pro Asp Gly Leu Val Lys Phe Ser Val 115 120 125	743
20	GAC GCA GAG ATA AGA GAG ATC TAT AAC GAG GAG TGC CCC GTC GTC ACT Asp Ala Glu Ile Arg Glu Ile Tyr Asn Glu Glu Cys Pro Val Val Thr 130 135 140	791
25	GAC GTG TCC GTC CCC CTC GAC GGC CGC CAG TGG AGC CTC TCG ATT TTC Asp Val Ser Val Pro Leu Asp Gly Arg Gln Trp Ser Leu Ser Ile Phe 145 150 155	839
30	TCC TTT CCG ATG TTC AGA ACC GCC TAC GTC GCC GTA GCG AAC GTC GAG Ser Phe Pro Met Phe Arg Thr Ala Tyr Val Ala Val Ala Asn Val Glu 160 165 170	887
35	AAC AAG GAG ATG TCG CTC GAC GTT GTC AAC GAC CTC ATC GAG TGG CTC Asn Lys Glu Met Ser Leu Asp Val Val Asn Asp Leu Ile Glu Trp Leu 175 180 185 190	935
40	AAC AAT CTC GCC GAC TGG CGT TAT GTC GTT GAC TCT GAA CAG TGG ATT Asn Asn Leu Ala Asp Trp Arg Tyr Val Val Asp Ser Glu Gln Trp Ile 195 200 205	983
45	AAC TTC ACC AAT GAC ACC ACG TAC TAC GTC CGC ATC CAC GTT CTA CGT Asn Phe Thr Asn Asp Thr Thr Tyr Tyr Val Arg Ile Arg Val Leu Arg 210 215 220	1031
50	CCA ACC TAC GAC GTT CCA GAC CCC ACA GAG GGC CTT GTT CGC ACA GTC Pro Thr Tyr Asp Val Pro Asp Pro Thr Glu Gly Leu Val Arg Thr Val 225 230 235	1079
55	TCA GAC TAC CGC CTC ACT TAT AAG GCG ATA ACA TGT GAA GCC AAC ATG Ser Asp Tyr Arg Leu Thr Tyr Lys Ala Ile Thr Cys Glu Ala Asn Met 240 245 250	1127
60	CCA ACA CTC GTC GAC CAA GGC TTT TGG ATC GGC GGC CAG TAC GCT CTC Pro Thr Leu Val Asp Gln Gly Phe Trp Ile Gly Gly Gln Tyr Ala Leu 255 260 265 270	1175
65	ACC CCG ACT AGC CTA CCG CAG TAC GAC GTC AGC GAG GCC TAC GCT CTG Thr Pro Thr Ser Leu Pro Gln Tyr Asp Val Ser Glu Ala Tyr Ala Leu 275 280 285	1223
70	CAC ACT TTG ACC TTC GCC AGA CCA TCC AGC GCC GCT GCA CTC GCG TTT His Thr Leu Thr Phe Ala Arg Pro Ser Ser Ala Ala Ala Leu Ala Phe 290 295 300	1271
75	GTG TGG GCA GGT TTG CCA CAG GGT GGC ACT GCG CCT GCA GGC ACT CCA Val Trp Ala Gly Leu Pro Gln Gly Gly Thr Ala Pro Ala Gly Thr Pro 305 310 315	1319
80	GCC TGG GAG CAG GCA TCC TCG GGT GGC TAC CTC ACC TGG CGC CAC AAC Ala Trp Glu Gln Ala Ser Ser Gly Gly Tyr Leu Thr Trp Arg His Asn 320 325 330	1367
85	GGT ACT ACT TTC CCA GCT GGC TCC GTT AGC TAC GTT CTC CCT GAG GGT Gly Thr Thr Phe Pro Ala Gly Ser Val Ser Tyr Val Leu Pro Glu Gly 335 340 345 350	1415
90	TTC GCC CTT GAG CGC TAC GAC CCG AAC GAC GGC TCT TGG ACC GAC TTC Phe Ala Leu Glu Arg Tyr Asp Pro Asn Asp Gly Ser Trp Thr Asp Phe 355 360 365	1463
95	GCT TCC GCA GGA GAC ACC GTC ACT TTC CGG CAG GTC GCC GTC GAC GAG Ala Ser Ala Gly Asp Thr Val Thr Phe Arg Gln Val Ala Val Asp Glu 370 375 380	1511
100	GTC GTT GTG ACC AAC AAC CCC GCC GGC GGC GGC AGC GCC CCC ACC TTC Val Val Val Thr Asn Asn Pro Ala Gly Gly Gly Ser Ala Pro Thr Phe 385 390 395	1559
105	ACC GTG AGA GTG CCC CCT TCA AAC GCT TAC ACC AAC ACC GTG TTT AGG Thr Val Arg Val Pro Pro Asn Ala Tyr Thr Asn Thr Val Phe Arg 400 405 410	1607

	AAC ACG CTC TTA GAG ACT CGA CCC TCC TCT CGT AGG CTC GAA CTC CCT Asn Thr Leu Leu Glu Thr Arg Pro Ser Ser Arg Arg Leu Glu Leu Pro 415 420 425 430	1655
5	ATG CCA CCT GCT GAC TTT GGA CAG ACG GTC GCC AAC AAC CCG AAG ATC Met Pro Pro Ala Asp Phe Gly Gln Thr Val Ala Asn Asn Pro Lys Ile 435 440 445	1703
10	GAG CAG TCG CTT CTT AAA GAA ACA CTT GGC TGC TAT TTG GTC CAC TCC Glu Gln Ser Leu Leu Lys Glu Thr Leu Gly Cys Tyr Leu Val His Ser 450 455 460	1751
15	AAA ATG CGA AAC CCC GTT TTC CAG CTC ACG CCA GCC AGC TCC TTT GGC Lys Met Arg Asn Pro Val Phe Gln Leu Thr Pro Ala Ser Ser Phe Gly 465 470 475	1799
20	GCC GTT TCC TTC AAC AAT CCG GGT TAT GAG CGC ACA CGC GAC CTC CCG Ala Val Ser Phe Asn Asn Pro Gly Tyr Glu Arg Thr Arg Asp Leu Pro 480 485 490	1847
	GAC TAC ACT GGC ATC CGT GAC TCA TTC GAC CAG AAC ATG TCC ACC GCT Asp Tyr Thr Gly Ile Arg Asp Ser Phe Asp Gln Asn Met Ser Thr Ala 495 500 505 510	1895
25	GTG GCC CAC TTC CGC TCA CTC TCC CAC TCC TGC AGT ATC GTC ACT AAG Val Ala His Phe Arg Ser Leu Ser His Ser Cys Ser Ile Val Thr Lys 515 520 525	1943
30	ACC TAC CAG GGT TGG GAA GGC GTC ACG AAC GTC AAC ACG CCT TTC GGC Thr Tyr Gln Gly Trp Glu Gly Val Thr Asn Val Asn Thr Pro Phe Gly 530 535 540	1991
35	CAA TTC GCG CAC GCG GGC CTC CTC AAG AAT GAG GAG ATC CTC TGC CTC Gln Phe Ala His Ala Gly Leu Lys Asn Glu Glu Ile Leu Cys Leu 545 550 555	2039
40	GCC GAC GAC CTG GCC ACC CGT CTC ACA GGT GTC TAC CCC GCC ACT GAC Ala Asp Asp Leu Ala Thr Arg Leu Thr Gly Val Tyr Pro Ala Thr Asp 560 565 570	2087
	AAC TTC GCG GCC GCC GTT TCT GCC TTC GCC GCG AAC ATG CTG TCC TCC Asn Phe Ala Ala Ala Val Ser Ala Phe Ala Ala Asn Met Leu Ser Ser 575 580 585 590	2135
45	GTG CTG AAG TCG GAG GCA ACG TCC TCC ATC ATC AAG TCC GTT GGC GAG Val Leu Lys Ser Glu Ala Thr Ser Ser Ile Ile Lys Ser Val Gly Glu 595 600 605	2183
50	ACT GCC GTC GGC GCG GCT CAG TCC GGC CTC GCG AAG CTA CCC GGA CTG Thr Ala Val Gly Ala Ala Gln Ser Gly Leu Ala Lys Leu Pro Gly Leu 610 615 620	2231
55	CTA ATG AGT GTA CCA GGG AAG ATT GCC GCG CGT GTC CGC GCG CGC CGA Leu Met Ser Val Pro Gly Lys Ile Ala Ala Arg Val Arg Ala Arg Arg 625 630 635	2279
60	GCG CGC CGC CGC GCC GCT CGT GCC AAT TAGTTTGCTC GTCCTGTTT Ala Arg Arg Arg Ala Ala Arg Ala Asn 640 645	2326
	CGCCGTTTCG TAAACGGCG TGTCGCCGA CATTACGGT ACCCTAAAGA CTCTGGTGAG TCCCCGTCGT TACACGACGG GTTCGCCGG GTTCGATTCC ATTCCCAAGC GGCAAGAAGG	2386 2446
65	ACGTAGTTAG CTCTGCGTCC CTCGGGATAC CA	2478
70	(2) INFORMATION FOR SEQ ID NO:50: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 647 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear	
75	(ii) MOLECULE TYPE: protein	
80	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50: Met Gly Asp Ala Gly Val Ala Ser Gln Arg Pro His Asn Arg Arg Gly 1 5 10 15	
85	Thr Arg Asn Val Arg Val Ser Ala Asn Thr Val Thr Val Asn Gly Arg 20 25 30	

	Arg	Asn	Gln	Arg	Arg	Arg	Thr	Gly	Arg	Gln	Val	Ser	Pro	Pro	Asp	Asn		
	35						40						45					
5	Phe	Thr	Ala	Ala	Ala	Gln	Asp	Leu	Ala	Gln	Ser	Leu	Asp	Ala	Asn	Thr		
	50						55			60								
	Val	Thr	Phe	Pro	Ala	Asn	Ile	Ser	Ser	Met	Pro	Glu	Phe	Arg	Asn	Trp		
10	65			70						75						80		
	Ala	Lys	Gly	Lys	Ile	Asp	Leu	Asp	Ser	Asp	Ser	Ile	Gly	Trp	Tyr	Phe		
				85						90						95		
15	Lys	Tyr	Leu	Asp	Pro	Ala	Gly	Ala	Thr	Glu	Ser	Ala	Arg	Ala	Val	Gly		
				100						105						110		
	Glu	Tyr	Ser	Lys	Ile	Pro	Asp	Gly	Leu	Val	Lys	Phe	Ser	Val	Asp	Ala		
				115			120						125					
20	Glu	Ile	Arg	Glu	Ile	Tyr	Asn	Glu	Glu	Cys	Pro	Val	Val	Thr	Asp	Val		
	130						135						140					
	Ser	Val	Pro	Leu	Asp	Gly	Arg	Gln	Trp	Ser	Leu	Ser	Ile	Phe	Ser	Phe		
25	145			150						155						160		
	Pro	Met	Phe	Arg	Thr	Ala	Tyr	Val	Ala	Val	Ala	Asn	Val	Glu	Asn	Lys		
				165						170						175		
30	Glu	Met	Ser	Leu	Asp	Val	Val	Asn	Asp	Leu	Ile	Glu	Trp	Leu	Asn	Asn		
	180									185						190		
	Leu	Ala	Asp	Trp	Arg	Tyr	Val	Val	Asp	Ser	Glu	Gln	Trp	Ile	Asn	Phe		
	195						200						205					
35	Thr	Asn	Asp	Thr	Thr	Tyr	Tyr	Val	Arg	Ile	Arg	Val	Leu	Arg	Pro	Thr		
	210						215						220					
	Tyr	Asp	Val	Pro	Asp	Pro	Thr	Glu	Gly	Leu	Val	Arg	Thr	Val	Ser	Asp		
40	225			230						235						240		
	Tyr	Arg	Leu	Thr	Tyr	Lys	Ala	Ile	Thr	Cys	Glu	Ala	Asn	Met	Pro	Thr		
				245						250						255		
45	Leu	Val	Asp	Gln	Gly	Phe	Trp	Ile	Gly	Gly	Gln	Tyr	Ala	Leu	Thr	Pro		
	260						265						270					
	Thr	Ser	Leu	Pro	Gln	Tyr	Asp	Val	Ser	Glu	Ala	Tyr	Ala	Leu	His	Thr		
	275						280						285					
50	Leu	Thr	Phe	Ala	Arg	Pro	Ser	Ser	Ala	Ala	Ala	Leu	Ala	Phe	Val	Trp		
	290						295						300					
	Ala	Gly	Leu	Pro	Gln	Gly	Gly	Thr	Ala	Pro	Ala	Gly	Thr	Pro	Ala	Trp		
55	305			310						315						320		
	Glu	Gln	Ala	Ser	Ser	Gly	Gly	Tyr	Leu	Thr	Trp	Arg	His	Asn	Gly	Thr		
				325						330						335		
60	Thr	Phe	Pro	Ala	Gly	Ser	Val	Ser	Tyr	Val	Leu	Pro	Glu	Gly	Phe	Ala		
	340									345						350		
	Leu	Glu	Arg	Tyr	Asp	Pro	Asn	Asp	Gly	Ser	Trp	Thr	Asp	Phe	Ala	Ser		
	355						360						365					
65	Ala	Gly	Asp	Thr	Val	Thr	Phe	Arg	Gln	Val	Ala	Val	Asp	Glu	Val	Val		
	370						375						380					
	Val	Thr	Asn	Asn	Pro	Ala	Gly	Gly	Gly	Ser	Ala	Pro	Thr	Phe	Thr	Val		
70	385			390						395						400		
	Arg	Val	Pro	Pro	Ser	Asn	Ala	Tyr	Thr	Asn	Thr	Val	Phe	Arg	Asn	Thr		

Ser Phe Asn Asn Pro Gly Tyr Glu Arg Thr Arg Asp Leu Pro Asp Tyr
 485 490 495
 5 Thr Gly Ile Arg Asp Ser Phe Asp Gln Asn Met Ser Thr Ala Val Ala
 500 505 510
 His Phe Arg Ser Leu Ser His Ser Cys Ser Ile Val Thr Lys Thr Tyr
 515 520 525
 10 Gln Gly Trp Glu Gly Val Thr Asn Val Asn Thr Pro Phe Gly Gln Phe
 530 535 540
 15 Ala His Ala Gly Leu Leu Lys Asn Glu Glu Ile Leu Cys Leu Ala Asp
 545 550 555 560
 Asp Leu Ala Thr Arg Leu Thr Gly Val Tyr Pro Ala Thr Asp Asn Phe
 565 570 575
 20 Ala Ala Ala Val Ser Ala Phe Ala Ala Asn Met Leu Ser Ser Val Leu
 580 585 590
 Lys Ser Glu Ala Thr Ser Ser Ile Ile Lys Ser Val Gly Glu Thr Ala
 595 600 605
 25 Val Gly Ala Ala Gln Ser Gly Leu Ala Lys Leu Pro Gly Leu Leu Met
 610 615 620
 Ser Val Pro Gly Lys Ile Ala Ala Arg Val Arg Ala Arg Arg Ala Arg
 625 630 635 640
 30 Arg Arg Ala Ala Arg Ala Asn
 645

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2479 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 283..2307

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

20	GTTCCTCTTT CTTTACCAAG TGTGGTAAAA TTAAACAAA GAAGAAAACC AGGACCGTAA	60
	CCCGGCCCTT ACACACCTCG AGTCGGTGAC CACCGGATTA TACGTGCCCC AOCACACGGC	120
25	GCCCTTTCCG ACCACTCTCG AGAGTCGTIG GGAGTTTCGT CCGTGACCAC CCGGTIGGCA	180
	GTCGACAGAC GCTTCCGSAC CACTAGAACC TCCTCGAGCG ACGCACACAC AGCACACACA	240
30	CCGCTTACG TGCACCTACG GCAGCGTTGA TAGCGCGGAT TT ATG AGC GAG CAC	294
	Met Ser Glu His	
	1	
	ACC ATC GCC CAC TCC ATC ACA TTA CCA CCC GGT TAC ACC CTT GCC CTA	342
	Thr Ile Ala His Ser Ile Thr Leu Pro Pro Gly Tyr Thr Leu Ala Leu	
	5 10 15 20	
35	ATA CCC CCT GAA CCT GAA GCA GGA TGG GAG ATG CTG GAG TGG CGT CAC	390
	Ile Pro Pro Glu Pro Glu Ala Gly Trp Glu Met Leu Glu Trp Arg His	
	25 30 35	
40	AGC GAC CTC ACA ACC GTC GCG GAA CCC GTA ACG TTC GGG TCA GCG CCA	438
	Ser Asp Leu Thr Thr Val Ala Glu Pro Val Thr Phe Gly Ser Ala Pro	
	40 45 50	
45	ACA CCG TCA CCG TCA ATG GTA GAA GAA ACC AAC GGC GTC GGA CCG GAA	486
	Thr Pro Ser Pro Ser Met Val Glu Thr Asn Gly Val Gly Pro Glu	
	55 60 65	
50	GGC AAG TTT CTC CCC CTG ACA ATT TCA CCG CTG CTG CAC AAG ACC TCG	534
	Gly Lys Phe Leu Pro Leu Thr Ile Ser Pro Leu Leu His Lys Thr Ser	
	70 75 80	
55	CGC AAA GCC TTG ACG CCA ACA CCG TCA CTT TCC CCC GCT AAC ATC TCT	582
	Arg Lys Ala Leu Thr Pro Thr Pro Ser Leu Ser Pro Ala Asn Ile Ser	
	85 90 95 100	
60	AGC ATG CCC GAA TTC CGG AAT TGG GCC AAG GGA AAG ATC GAC CTC GAC	630
	Ser Met Pro Glu Phe Arg Asn Trp Ala Lys Gly Lys Ile Asp Leu Asp	
	105 110 115	
65	TCC GAT TCC ATC GGC TGG TAC TTC AAG TAC CTT GAC CCA GCG GGT GCT	678
	Ser Asp Ser Ile Gly Trp Tyr Phe Lys Tyr Leu Asp Pro Ala Gly Ala	
	120 125 130	
70	ACA GAG TCT GCG CGC GCC GTC GGC GAG TAC TCG AAG ATC CCT GAC GGC	726
	Thr Glu Ser Ala Arg Ala Val Gly Glu Tyr Ser Lys Ile Pro Asp Gly	
	135 140 145	
75	CTC GTC AAG TTC TCC GTC GAC GCA GAG ATA AGA GAG ATC TAT AAC GAG	774
	Leu Val Lys Phe Ser Val Asp Ala Glu Ile Arg Glu Ile Tyr Asn Glu	
	150 155 160	
80	GAG TGC CCC GTC GTC ACT GAC GTG TCC GTC CCC CTC GAC GGC CGC CAG	822
	Glu Cys Pro Val Val Thr Asp Val Ser Val Pro Leu Asp Gly Arg Gln	
	165 170 175 180	
85	TGG AGC CTC TCG ATT TTC TCC TTT CCG ATG TTC AGA ACC GCC TAC GTC	870
	Trp Ser Leu Ser Ile Phe Ser Phe Pro Met Phe Arg Thr Ala Tyr Val	
	185 190 195	
85	GCC GTA GCG AAC GTC GAG AAC AAG GAG ATG TCG CTC GAC GTT GTC AAC	918
	Ala Val Ala Asn Val Glu Asn Lys Glu Met Ser Leu Asp Val Val Asn	
	200 205 210	
85	GAC CTC ATC GAG TGG CTC AAC AAT CTC GCC GAC TGG CGT TAT GTC GTT	966
	Asp Leu Ile Glu Trp Leu Asn Asn Leu Ala Asp Trp Arg Tyr Val Val	

099126 11001
 11002 11003
 11004 11005
 11006 11007
 11008 11009
 11010 11011
 11012 11013
 11014 11015
 11016 11017
 11018 11019
 11020 11021
 11022 11023
 11024 11025
 11026 11027
 11028 11029
 11030 11031
 11032 11033
 11034 11035
 11036 11037
 11038 11039
 11040 11041
 11042 11043
 11044 11045
 11046 11047
 11048 11049
 11050 11051
 11052 11053
 11054 11055
 11056 11057
 11058 11059
 11060 11061
 11062 11063
 11064 11065
 11066 11067
 11068 11069
 11070 11071
 11072 11073
 11074 11075
 11076 11077
 11078 11079
 11080 11081
 11082 11083
 11084 11085
 11086 11087
 11088 11089
 11090 11091
 11092 11093
 11094 11095
 11096 11097
 11098 11099
 11100 11101
 11102 11103
 11104 11105
 11106 11107
 11108 11109
 11110 11111
 11112 11113
 11114 11115
 11116 11117
 11118 11119
 11120 11121
 11122 11123
 11124 11125
 11126 11127
 11128 11129
 11130 11131
 11132 11133
 11134 11135
 11136 11137
 11138 11139
 11140 11141
 11142 11143
 11144 11145
 11146 11147
 11148 11149
 11150 11151
 11152 11153
 11154 11155
 11156 11157
 11158 11159
 11160 11161
 11162 11163
 11164 11165
 11166 11167
 11168 11169
 11170 11171
 11172 11173
 11174 11175
 11176 11177
 11178 11179
 11180 11181
 11182 11183
 11184 11185
 11186 11187
 11188 11189
 11190 11191
 11192 11193
 11194 11195
 11196 11197
 11198 11199
 11200 11201
 11202 11203
 11204 11205
 11206 11207
 11208 11209
 11210 11211
 11212 11213
 11214 11215
 11216 11217
 11218 11219
 11220 11221
 11222 11223
 11224 11225
 11226 11227
 11228 11229
 11230 11231
 11232 11233
 11234 11235
 11236 11237
 11238 11239
 11240 11241
 11242 11243
 11244 11245
 11246 11247
 11248 11249
 11250 11251
 11252 11253
 11254 11255
 11256 11257
 11258 11259
 11260 11261
 11262 11263
 11264 11265
 11266 11267
 11268 11269
 11270 11271
 11272 11273
 11274 11275
 11276 11277
 11278 11279
 11280 11281
 11282 11283
 11284 11285
 11286 11287
 11288 11289
 11290 11291
 11292 11293
 11294 11295
 11296 11297
 11298 11299
 11300 11301
 11302 11303
 11304 11305
 11306 11307
 11308 11309
 11310 11311
 11312 11313
 11314 11315
 11316 11317
 11318 11319
 11320 11321
 11322 11323
 11324 11325
 11326 11327
 11328 11329
 11330 11331
 11332 11333
 11334 11335
 11336 11337
 11338 11339
 11340 11341
 11342 11343
 11344 11345
 11346 11347
 11348 11349
 11350 11351
 11352 11353
 11354 11355
 11356 11357
 11358 11359
 11360 11361
 11362 11363
 11364 11365
 11366 11367
 11368 11369
 11370 11371
 11372 11373
 11374 11375
 11376 11377
 11378 11379
 11380 11381
 11382 11383
 11384 11385
 11386 11387
 11388 11389
 11390 11391
 11392 11393
 11394 11395
 11396 11397
 11398 11399
 11400 11401
 11402 11403
 11404 11405
 11406 11407
 11408 11409
 11410 11411
 11412 11413
 11414 11415
 11416 11417
 11418 11419
 11420 11421
 11422 11423
 11424 11425
 11426 11427
 11428 11429
 11430 11431
 11432 11433
 11434 11435
 11436 11437
 11438 11439
 11440 11441
 11442 11443
 11444 11445
 11446 11447
 11448 11449
 11450 11451
 11452 11453
 11454 11455
 11456 11457
 11458 11459
 11460 11461
 11462 11463
 11464 11465
 11466 11467
 11468 11469
 11470 11471
 11472 11473
 11474 11475
 11476 11477
 11478 11479
 11480 11481
 11482 11483
 11484 11485
 11486 11487
 11488 11489
 11490 11491
 11492 11493
 11494 11495
 11496 11497
 11498 11499
 11500 11501
 11502 11503
 11504 11505
 11506 11507
 11508 11509
 11510 11511
 11512 11513
 11514 11515
 11516 11517
 11518 11519
 11520 11521
 11522 11523
 11524 11525
 11526 11527
 11528 11529
 11530 11531
 11532 11533
 11534 11535
 11536 11537
 11538 11539
 11540 11541
 11542 11543
 11544 11545
 11546 11547
 11548 11549
 11550 11551
 11552 11553
 11554 11555
 11556 11557
 11558 11559
 11560 11561
 11562 11563
 11564 11565
 11566 11567
 11568 11569
 11570 11571
 11572 11573
 11574 11575
 11576 11577
 11578 11579
 11580 11581
 11582 11583
 11584 11585
 11586 11587
 11588 11589
 11590 11591
 11592 11593
 11594 11595
 11596 11597
 11598 11599
 11600 11601
 11602 11603
 11604 11605
 11606 11607
 11608 11609
 11610 11611
 11612 11613
 11614 11615
 11616 11617
 11618 11619
 11620 11621
 11622 11623
 11624 11625
 11626 11627
 11628 11629
 11630 11631
 11632 11633
 11634 11635
 11636 11637
 11638 11639
 11640 11641
 11642 11643
 11644 11645
 11646 11647
 11648 11649
 11650 11651
 11652 11653
 11654 11655
 11656 11657
 11658 11659
 11660 11661
 11662 11663
 11664 11665
 11666 11667
 11668 11669
 11670 11671
 11672 11673
 11674 11675
 11676 11677
 11678 11679
 11680 11681
 11682 11683
 11684 11685
 11686 11687
 11688 11689
 11690 11691
 11692 11693
 11694 11695
 11696 11697
 11698 11699
 11700 11701
 11702 11703
 11704 11705
 11706 11707
 11708 11709
 11710 11711
 11712 11713
 11714 11715
 11716 11717
 11718 11719
 11720 11721
 11722 11723
 11724 11725
 11726 11727
 11728 11729
 11730 11731
 11732 11733
 11734 11735
 11736 11737
 11738 11739
 11740 11741
 11742 11743
 11744 11745
 11746 11747
 11748 11749
 11750 11751
 11752 11753
 11754 11755
 11756 11757
 11758 11759
 11760 11761
 11762 11763
 11764 11765
 11766 11767
 11768 11769
 11770 11771
 11772 11773
 11774 11775
 11776 11777
 11778 11779
 11780 11781
 11782 11783
 11784 11785
 11786 11787
 11788 11789
 11790 11791
 11792 11793
 11794 11795
 11796 11797
 11798 11799
 11800 11801
 11802 11803
 11804 11805
 11806 11807
 11808 11809
 11810 11811
 11812 11813
 11814 11815
 11816 11817
 11818 11819
 11820 11821
 11822 11823
 11824 11825
 11826 11827
 11828 11829
 11830 11831
 11832 11833
 11834 11835
 11836 11837
 11838 11839
 11840 11841
 11842 11843
 11844 11845
 11846 11847
 11848 11849
 11850 11851
 11852 11853
 11854 11855
 11856 11857
 11858 11859
 11860 11861
 11862 11863
 11864 11865
 11866 11867
 11868 11869
 11870 11871
 11872 11873
 11874 11875
 11876 11877
 11878 11879
 11880 11881
 11882 11883
 11884 11885
 11886 11887
 11888 11889
 11890 11891
 11892 11893
 11894 11895
 11896 11897
 11898 11899
 11900 11901
 11902 11903
 11904 11905
 11906 11907
 11908 11909
 11910 11911
 11912 11913
 11914 11915
 11916 11917
 11918 11919
 11920 11921
 11922 11923
 11924 11925
 11926 11927
 11928 11929
 11930 11931
 11932 11933
 11934 11935
 11936 11937
 11938 11939
 11940 11941
 11942 11943
 11944 11945
 11946 11947
 11948 11949
 11950 11951
 11952 11953
 11954 11955
 11956 11957
 11958 11959
 11960 11961
 11962 11963
 11964 11965
 11966 11967
 11968 11969
 11970 11971
 11972 11973
 11974 11975
 11976 11977
 11978 11979
 11980 11981
 11982 11983
 11984 11985
 11986 11987
 11988 11989
 11990 11991
 11992 11993
 11994 11995
 11996 11997
 11998 11999
 12000 12001
 12002 12003
 12004 12005
 12006 12007
 12008 12009
 12010 12011
 12012 12013
 12014 12015
 12016 12017
 12018 12019
 12020 12021
 12022 12023
 12024 12025
 12026 12027
 12028 12029
 12030 12031
 12032 12033
 12034 12035
 12036 12037
 12038 12039
 12040 12041
 12042 12043
 12044 12045
 12046 12047
 12048 12049
 12050 12051
 12052 12053
 12054 12055
 12056 12057
 12058 12059
 12060 12061
 12062 12063
 12064 12065
 12066 12067
 12068 12069
 12070 12071
 12072 12073
 12074 12075
 12076 12077
 12078 12079
 12080 12081
 12082 12083
 12084 12085
 12086 12087
 12088 12089
 12090 12091
 12092 12093
 12094 12095
 12096 12097
 12098 12099
 12100 12101
 12102 12103
 12104 12105
 12106 12107
 12108 12109
 12110 12111
 12112 12113
 12114 12115
 12116 12117
 12118 12119
 12120 12121
 12122 12123
 12124 12125
 12126 12127
 12128 12129
 12130 12131
 12132 12133
 12134 12135
 12136 12137
 12138 12139
 12140 12141
 12142 12143
 12144 12145
 12146 12147
 12148 12149
 12150 12151
 12152 12153
 12154 12155
 12156 12157
 12158 12159
 12160 12161
 12162 12163
 12164 12165
 12166 12167
 12168 12169
 12170 12171
 12172 12173
 12174 12175
 12176 12177
 12178 12179
 12180 12181
 12182 12183
 12184 12185
 12186 12187
 12188 12189
 12190 12191
 12192 12193
 12194 12195
 12196 12197
 12198 12199
 12200 12201
 12202 12203
 12204 12205
 12206 12207
 12208 12209
 12210 12211
 12212 12213
 12214 12215
 12216 12217
 12218 12219
 12220 12221
 12222 12223
 12224 12225
 12226 12227
 12228 12229
 12230 12231
 12232 12233
 12234 12235
 12236 12237
 12238 12239

5

5	GGC CTT GTT CGC ACA GTC TCA GAC TAC CGC CTC ACT TAT AAG GCG ATA Gly Leu Val Arg Thr 265 Val Ser Asp Tyr Arg Leu Thr Tyr Lys Ala Ile 275	1110
10	ACA TGT GAA GCC AAC ATG CCA ACA CTC GTC GAC CAA GGC TTT TGG ATC Thr Cys Glu Ala Asn Met Pro Thr 285 Leu Val Asp Gln Gly Phe Trp Ile 290	1158
15	GGC GGC CAG TAC GCT CTC ACC CCG ACT AGC CTA CCG CAG TAC GAC GTC Gly Gly Gln Tyr Ala Leu Thr Pro Thr Ser Leu Pro Gln Tyr Asp Val 305	1206
20	AGC GAG GCC TAC GCT CTG CAC ACT TTG ACC TTC GCC AGA CCA TCC AGC Ser Glu Ala Tyr Ala Leu His 315 Thr Leu Thr Phe Ala Arg Pro Ser Ser 320	1254
25	GCC GCT GCA CTC GCG TTT GTG TGG GCA GGT TTG CCA CAG GGT GGC ACT Ala Ala Ala Leu Ala Phe Val Trp Ala Gly Leu Pro Gln Gly Gly Thr 330 335 340	1302
30	GCG CCT GCA GGC ACT CCA GCC TGG GAG CAG GCA TCC TCG GGT GGC TAC Ala Pro Ala Gly Thr 345 Pro Ala Trp Glu Gln Ala Ser Ser Gly Gly Tyr 355	1350
35	CTC ACC TGG CGC CAC AAC GGT ACT ACT TTC CCA GCT GGC TCC GTT AGC Leu Thr Trp Arg His Asn Gly Thr 360 Thr Phe Pro Ala Gly Ser Val Ser 370	1398
40	TAC GTT CTC CCT GAG GGT TTC GCC CTT GAG CGC IAC GAC CCG AAC GAC Tyr Val Leu Pro Glu Gly Phe Ala Leu Glu Arg Tyr Asp Pro Asn Asp 375 380 385	1446
45	GGC ICT TGG ACC GAC TTC GCT TCC GCA GGA GAC ACC GTC ACT TTC CGG Gly Ser Trp Thr Asp Phe Ala Ser Ala Gly Asp Thr Val Thr Phe Arg 390 395 400	1494
50	CAG GTC GCC GTC GAC GAG GTC GTT GTG ACC AAC AAC CCC GCC GGC GGC Gln Val Ala Val Asp Glu Val Val Val Thr Asn Asn Pro Ala Gly Gly 405 410 415 420	1542
55	GGC AGC GCC CCC ACC TTC ACC GTG AGA GTG CCC CCT TCA AAC GCT TAC Gly Ser Ala Pro Thr Phe Thr Val Arg Val Pro Pro Ser Asn Ala Tyr 425 430 435	1590
60	ACC AAC ACC GTG TTT AGG AAC ACG CTC TTA GAG ACT CGA CCC TCC TCT Thr Asn Thr Val Phe Arg Asn Thr Leu Leu Glu Thr Arg Pro Ser Ser 440 445 450	1638
65	CGT AGG CTC GAA CTC CCT ATG CCA CCT GCT GAC TTT GGA CAG ACG GTC Arg Arg Leu Glu Leu Pro Met Pro Pro Ala Asp Phe Gly Gln Thr Val 455 460 465	1686
70	GCC AAC AAC CCG AAG ATC GAG CAG TCG CTT CTT AAA GAA ACA CTT GGC Ala Asn Asn Pro Lys Ile Glu Gln Ser Leu Leu Lys Glu Thr Leu Gly 470 475 480	1734
75	TGC TAT TTG GTC CAC TCC AAA ATG CGA AAC CCC GTT TTC CAG CTC ACG Cys Tyr Leu Val His Ser Lys Met Arg Asn Pro Val Phe Gln Leu Thr 485 490 495 500	1782
80	CCA GCC AGC TCC TTT GGC GCC GTT TCC TTC AAC AAT CCG GGT TAT GAG Pro Ala Ser Ser Phe Gly Ala Val Ser Phe Asn Asn Pro Gly Tyr Glu 505 510 515	1830
85	CGC ACA CGC GAC CTC CCG GAC TAC ACT GGC ATC CGT GAC TCA TTC GAC Arg Thr Arg Asp Leu Pro Asp Tyr Thr Gly Ile Arg Asp Ser Phe Asp 520 525 530	1878
90	CAG AAC ATG TCC ACC GCT GTG GCC CAC TTC CGC TCA CTC TCC CAC TCC Gln Asn Met Ser Thr Ala Val Ala His Phe Arg Ser Leu Ser His Ser 535 540 545	1926
95	TGC AGT ATC GTC ACT AAG ACC TAC CAG GGT TGG GAA GGC GTC ACG AAC Cys Ser Ile Val Thr Lys Thr Tyr Gln Gly Trp Glu Gly Val Thr Asn 550 555 560	1974
100	GTC AAC ACG CCT TTC GGC CAA TTC GCG CAC GCG GGC CTC CTC AAG AAT Val Asn Thr Pro Phe Gly Gln Phe Ala His Ala Gly Leu Leu Lys Asn 565 570 575 580	2022
105	GAG GAG ATC CTC TGC CTC GCC GAC GAC CTG GCC ACC CGT CTC ACA GGT Glu Glu Ile Leu Cys Leu Ala Asp Asp Leu Ala Thr Arg Leu Thr Gly 585 590 595	2070

	GTC TAC CCC GCC ACT GAC AAC TTC GCG GCC GCC GTT TCT GCC TTC GCC	2118
	Val Tyr Pro Ala Thr Asp Asn Phe Ala Ala Val Ser Ala Phe Ala	
	600 605 610	
5	GCG AAC ATG CTG TCC TCC GTG CTG AAG TCG GAG GCA ACG TCC TCC ATC	2166
	Ala Asn Met Leu Ser Ser Val Leu Lys Ser Glu Ala Thr Ser Ser Ile	
	615 620 625	
10	ATC AAG TCC GTT GGC GAG ACT GCC GTC GGC GCG GCT CAG TCC GGC CTC	2214
	Ile Lys Ser Val Gly Glu Thr Ala Val Gly Ala Ala Gln Ser Gly Leu	
	630 635 640	
	GCG AAG CTA CCC GGA CTG CTA ATG AGT GTA CCA GGG AAG ATT GCC GCG	2262
15	Ala Lys Leu Pro Gly Leu Leu Met Ser Val Pro Gly Lys Ile Ala Ala	
	645 650 655 660	
	CGT GTC CGC GCG CGC CGA GCG CGC CGC GCG GCT CGT GCC AAT	2307
	Arg Val Arg Ala Arg Arg Ala Arg Arg Arg Ala Ala Arg Ala Asn	
	665 670 675	
20	TAGTTTGCTC GCTCCTGTITT CGCCGTTTCG TAAACGGCG TGGTCCCGCA CATTACGCGT	2367
	ACCCCTAAAGA CTCTGGTGAG TCCCGTTCGT TACACGACGG GTCGCGCGG GTTCGATTCC	2427
25	ATTCCTCAAGC GGCAAGAAGG ACGTAGTTAG CTCTGCGTCC CTCGGGATAC CA	2479
30	(2) INFORMATION FOR SEQ ID NO:52:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 675 amino acids	
	(B) TYPE: amino acid	
	(D) TOPOLOGY: linear	
35	(ii) MOLECULE TYPE: protein	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:	
40	Met Ser Glu His Thr Ile Ala His Ser Ile Thr Leu Pro Pro Gly Tyr	
	1 5 10 15	
45	Thr Leu Ala Leu Ile Pro Pro Glu Pro Glu Ala Gly Trp Glu Met Leu	
	20 25 30	
	Glu Trp Arg His Ser Asp Leu Thr Thr Val Ala Glu Pro Val Thr Phe	
	35 40 45	
50	Gly Ser Ala Pro Thr Pro Ser Pro Ser Met Val Glu Glu Thr Asn Gly	
	50 55 60	
	Val Gly Pro Glu Gly Lys Phe Leu Pro Leu Thr Ile Ser Pro Leu Leu	
55	65 70 75 80	
	His Lys Thr Ser Arg Lys Ala Leu Thr Pro Thr Pro Ser Leu Ser Pro	
	85 90 95	
60	Ala Asn Ile Ser Ser Met Pro Glu Phe Arg Asn Trp Ala Lys Gly Lys	
	100 105 110	
	Ile Asp Leu Asp Ser Asp Ser Ile Gly Trp Tyr Phe Lys Tyr Leu Asp	
	115 120 125	
65	Pro Ala Gly Ala Thr Glu Ser Ala Arg Ala Val Gly Glu Tyr Ser Lys	
	130 135 140	
	Ile Pro Asp Gly Leu Val Lys Phe Ser Val Asp Ala Glu Ile Arg Glu	
70	145 150 155 160	
	Ile Tyr Asn Glu Glu Cys Pro Val Val Thr Asp Val Ser Val Pro Leu	
	165 170 175	
75	Asp Gly Arg Gln Trp Ser Leu Ser Ile Phe Ser Phe Pro Met Phe Arg	
	180 185 190	
	Thr Ala Tyr Val Ala Val Ala Asn Val Glu Asn Lys Glu Met Ser Leu	
	195 200 205	
80	Asp Val Val Asn Asp Leu Ile Glu Trp Leu Asn Asn Leu Ala Asp Trp	
	210 215 220	
85	Arg Tyr Val Val Asp Ser Glu Gln Trp Ile Asn Phe Thr Asn Asp Thr	
	225 230 235 240	

5	Thr	Tyr	Tyr	Val	Arg	Ile	Arg	Val	Leu	Arg	Pro	Thr	Tyr	Asp	Val	Pro
	245															
	Asp	Pro	Thr	Glu	Gly	Leu	Val	Arg	Thr	Val	Ser	Asp	Tyr	Arg	Leu	Thr
10	Tyr	Lys	Ala	Ile	Thr	Cys	Glu	Ala	Asn	Met	Pro	Thr	Leu	Val	Asp	Gln
	275															
	Gly	Phe	Trp	Ile	Gly	Gly	Gln	Tyr	Ala	Leu	Thr	Pro	Thr	Ser	Leu	Pro
15	Gln	Tyr	Asp	Val	Ser	Glu	Ala	Tyr	Ala	Leu	His	Thr	Leu	Thr	Phe	Ala
	305															
	Arg	Pro	Ser	Ser	Ala	Ala	Ala	Leu	Ala	Phe	Val	Trp	Ala	Gly	Leu	Pro
20	Gln	Gly	Gly	Thr	Ala	Pro	Ala	Gly	Thr	Pro	Ala	Trp	Glu	Gln	Ala	Ser
	340															
	Ser	Gly	Gly	Tyr	Leu	Thr	Trp	Arg	His	Asn	Gly	Thr	Thr	Phe	Pro	Ala
25	Gly	Ser	Val	Ser	Tyr	Val	Leu	Pro	Glu	Gly	Phe	Ala	Leu	Glu	Arg	Tyr
	370															
	Asp	Pro	Asn	Asp	Gly	Ser	Trp	Thr	Asp	Phe	Ala	Ser	Ala	Gly	Asp	Thr
30	Val	Thr	Phe	Arg	Gln	Val	Ala	Val	Asp	Glu	Val	Val	Val	Thr	Asn	Asn
	405															
	Pro	Ala	Gly	Gly	Gly	Ser	Ala	Pro	Thr	Phe	Trp	Val	Arg	Val	Pro	Pro
35	Ser	Asn	Ala	Tyr	Thr	Asn	Thr	Val	Phe	Arg	Asn	Thr	Leu	Leu	Glu	Thr
	435															
	Arg	Pro	Ser	Ser	Arg	Arg	Leu	Glu	Leu	Pro	Met	Pro	Pro	Ala	Asp	Phe
40	Gly	Gln	Thr	Val	Ala	Asn	Asn	Pro	Lys	Ile	Glu	Gln	Ser	Leu	Leu	Lys
	465															
	Glu	Thr	Leu	Gly	Cys	Tyr	Leu	Val	His	Ser	Lys	Met	Arg	Asn	Pro	Val
50	Phe	Gln	Leu	Thr	Pro	Ala	Ser	Ser	Phe	Gly	Ala	Val	Ser	Phe	Asn	Asn
	500															
	Pro	Gly	Tyr	Glu	Arg	Thr	Arg	Asp	Leu	Pro	Asp	Tyr	Thr	Gly	Ile	Arg
55	Asp	Ser	Phe	Asp	Gln	Asn	Met	Ser	Thr	Ala	Val	Ala	His	Phe	Arg	Ser
	530															
	Leu	Ser	His	Ser	Cys	Ser	Ile	Val	Thr	Lys	Thr	Tyr	Gln	Gly	Trp	Glu
60	Gly	Val	Thr	Asn	Val	Asn	Thr	Pro	Phe	Gly	Gln	Phe	Ala	His	Ala	Gly
	565															
	Leu	Leu	Lys	Asn	Glu	Glu	Ile	Leu	Cys	Leu	Ala	Asp	Asp	Leu	Ala	Thr
65	Arg	Leu	Thr	Gly	Val	Tyr	Pro	Ala	Thr	Asp	Asn	Phe	Ala	Ala	Ala	Val
	595															
	Ser	Ala	Phe	Ala	Ala	Asn	Met	Leu	Ser	Ser	Val	Leu	Lys	Ser	Glu	Ala
70	Thr	Ser	Ser	Ile	Ile	Lys	Ser	Val	Gly	Glu	Thr	Ala	Val	Gly	Ala	Ala
	625															
	Gln	Ser	Gly	Leu	Ala	Lys	Leu	Pro	Gly	Leu	Leu	Met	Ser	Val	Pro	Gly
75	Lys	Ile	Ala	Ala	Arg	Val	Arg	Ala	Arg	Arg	Ala	Arg	Arg	Arg	Ala	Ala
	660															
	Arg	Ala	Asn													
80	675															
85																

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

GGGGATCCAC AGTTCCTGCCT CCCCCGGACG GTAAATATAG GGGGAACCATG GTCTAGAGG

59